



UNIVERSITY OF  
BIRMINGHAM

# **SUPPORTING STUDENTS' CONFIDENCE JUDGEMENT THROUGH VISUALISING ALIGNMENT IN OPEN LEARNER MODELS**

**Lamiya Said Salim Al-Shanfari**

A thesis submitted to the University of Birmingham for the degree of  
DOCTOR OF PHILOSOPHY

School of Electronic, Electrical and  
Systems Engineering  
College of Engineering and Physical Sciences  
University of Birmingham  
June 2019

UNIVERSITY OF  
BIRMINGHAM

**University of Birmingham Research Archive**

**e-theses repository**

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

# ABSTRACT

Supporting students' knowledge monitoring skills, a component of metacognition, can help students regulate their own learning. This thesis investigates the alignment of learners' confidence in their knowledge with a computer's assessment of their knowledge, visualised using an Open Learner Model (OLM). The research explored students' preferred method for visualising inconsistent data (e.g. misalignment) in an OLM, and the ways in which visualising alignment can influence student interaction with the computer.

The thesis demonstrates that visualising alignment in Open Learner Models significantly increases students' confidence compared to a control condition. In particular, visualising alignment benefited low-achieving students, in terms of knowledge monitoring and this was associated with improvements in their performance. Students showed a preference towards the visualisations that provides an overview of the information (i.e. opacity) rather than ones, which provide detailed information. Graphical representation is shown to be more beneficial in motivating students to interact with the system than text-based representation of the same information in the context of representing the alignment within OLMs.

# ACKNOWLEDGMENTS

In the name of Allah, the Most Gracious and the Most Merciful.

I would like to express my honest gratitude to my supervisor Prof. Chris Baber for his aspiring guidance, support and constructive criticism. Special thanks to my supervisor Dr. Susan Bull for her valuable support and patience in the first year of my journey of the PhD study.

I am extremely thankful to the co-author Carrie Demmans Epp for her collaboration in the research papers I produced through the period of my PhD study. Her never ending support, inspirations and guidance and the assistance of my supervisors encouraged me to publish my work in top conferences and journals in the field such as Artificial Intelligent in Education, Learning Analytics and Knowledge and User Modelling and User Adapted Interaction.

I would like to express my appreciation for every person who supported me to complete this thesis. I would like to thank my beloved family especially to my husband Ahmed who supported me since day one and his trust that I can achieve it and my daughters Ayah, Muna and Tafool. Thank you to my parents, my sisters and brothers for all their support.

# Table of Contents

<b>Chapter 1 : INTRODUCTION .....</b>	<b>1</b>
1.1 Research Questions: .....	5
1.2 Metacognition .....	5
1.3 Measures of Metacognitive Judgement .....	6
1.4 Can Students Accurately Assess their own Knowledge? .....	10
1.5 Study Objectives and Approach.....	11
1.6 Chronology of the work .....	13
1.7 Contributions.....	15
1.8 Thesis Summary.....	17
1.9 Outline of Chapters .....	20
<b>Chapter 2 : Open Learner Models .....</b>	<b>23</b>
2.1 Intelligent Tutoring Systems.....	23
2.2 Open Learner Models .....	24
2.3 Supporting Students' Metacognitive Skills in Open Learner Models.....	28
2.4 Aligning Multiple Belief Sources in Open Learner Models .....	33
2.5 Comparing Visualisations within OLMs.....	37

## **Chapter 3 : VISUALISING ALIGNMENT WITHIN OPEN LEARNER MODELS**

<b>(OLMS).....</b>	<b>40</b>
3.1 Information Visualisation.....	40
3.2 Different Methods to Visualise Alignment in OLMs.....	45
3.3 System design and architecture (The OLMlets-OLM System).....	55
3.3.1 The OLMlets interface.....	56
3.3.2 OLMlets Base system.....	58
3.3.3 Modelling alignment.....	63
3.3.4. Visualising alignment.....	65
3.3.5 A pilot study for an alignment representation decision.....	68
3.3.6 The impact of students' answers on the learner model.....	71
3.3.7 Representation of alignment using text-based.....	74

## **Chapter 4 : PREFERRED METHOD OF VISUALISING ALIGNMENT IN OLM**

<b>(EXPLORATORY STUDY).....</b>	<b>78</b>
4.1 Overview.....	78
4.2 Visual variables applied in OLM.....	80
4.3 Visual variables selected for this study.....	81
4.4 Designing the visual variables into OLM visualisation (skill meter).....	83
4.5 Method.....	85
4.5.1 Study design.....	85

4.5.2 Participants .....	86
4.5.3 Study procedures .....	86
4.5.4 Analysis .....	87
4.5.5 Data collection .....	89
4.6 Results .....	89
4.6.1 Thurstone's paired comparison method .....	89
4.6.2 Statistical analysis .....	93
4.7 Discussion .....	96

## **Chapter 5 : EVALUATING THE EFFECT OF VISUALISING ALIGNMENT ON**

### **STUDENTS' KNOWLEDGE MONITORING SKILLS .....98**

5.1 Overview .....	98
5.2 Visualising alignment in OLM .....	99
5.3 Method .....	103
5.3.1 The OLMlets OLM .....	103
5.3.2 Study design .....	106
5.3.3 Study procedures .....	107
5.3.4 Instruments .....	108
5.3.5 Study analysis .....	109
5.3.6 Participants .....	110

5.4 Results.....	111
5.4.1 Normality of data.....	111
5.4.2 Student awareness from the pre- and post-questionnaire .....	112
5.4.3 Pre- and post-test of student knowledge.....	122
5.4.4 Pre- and post-test of student confidence .....	124
5.4.5 System use .....	128
5.4.6 Relationship of system use to confidence .....	132
5.5 Discussion .....	132

## **Chapter 6 : EVALUATING THE EFFECT OF VISUALISING ALIGNMENT ON**

### **STUDENTS' BIAS OF CONFIDENCE JUDGEMENT ..... 135**

6.1 Overview.....	135
6.2 Introduction .....	136
6.3 Method.....	139
6.3.1 Participants .....	139
6.3.2 Study design .....	140
6.3.3 System training .....	143
6.3.4 Design of the questions integrated into OLMlets (Java Programming) ....	143
6.3.5 Data collection .....	147
6.3.6 Measurement and quantitative analysis procedures .....	148



6.3.7 Qualitative analysis procedures.....	150
6.4 Results.....	151
6.4.1 Normality of data.....	151
6.4.2 Students' bias of confidence judgement .....	152
6.4.3 Students' bias of confidence judgement over time: potential practice effects .....	157
6.4.4 System usage by condition .....	160
6.5 Discussion .....	167
6.5.1 The effect of OLMlets and visualising alignment on students' bias of confidence judgement.....	167

**Chapter 7 : EVALUATING THE EFFECT OF THE REPRESENTATION OF  
ALIGNMENT ON STUDENTS' CONFIDENCE JUDGEMENT .....171**

7.1 Overview.....	171
7.2 Introduction .....	172
7.3 Method.....	173
7.3.1 Participants .....	173
7.3.2 Study design .....	174
7.3.3 Design of the questions integrated to OLMlets .....	177
7.3.4 Data collection .....	179
7.3.5 Measures .....	179

7.3.6 Quantitative analysis .....	180
7.3.7 Qualitative analysis .....	181
7.4 Results .....	181
7.4.1 Normality of data .....	181
7.4.2 Students' confidence judgement: visual vs. textual OLMs .....	182
7.4.3 System usage .....	186
7.5 Discussion .....	190
7.5.1 Effect of the OLM representation on students' bias of confidence judgement .....	191
<b>Chapter 8 : CONCLUSIONS AND FUTURE WORK.....</b>	<b>194</b>
8.1 Context .....	194
8.2 Findings .....	195
8.3 Contribution .....	199
8.4 Limitations and Future Work .....	201
<b>List of References .....</b>	<b>203</b>
<b>Appendix 1: Paired Comparison Survey (Chapter 4) .....</b>	<b>221</b>
<b>Appendix 2: Paired Comparison Survey: Student data (SPSS file) .....</b>	<b>225</b>
<b>Appendix 3: Pre-questionnaire (Chapter 5).....</b>	<b>227</b>
<b>Appendix 4: Pre-questionnaire students' responses (Chapter 5).....</b>	<b>229</b>

<b>Appendix 5: Post-questionnaire control group (Chapter 5) .....</b>	<b>232</b>
<b>Appendix 6: Post-questionnaire combined group (Chapter 5).....</b>	<b>234</b>
<b>Appendix 7: Post-questionnaire expandable group (Chapter 5).....</b>	<b>236</b>
<b>Appendix 8: Post-Questionnaire students responses (Chapter 5) .....</b>	<b>238</b>
<b>Appendix 9: Pre-test (Chapter 5).....</b>	<b>243</b>
<b>Appendix 10: Post-test (Chapter 5).....</b>	<b>247</b>
<b>Appendix 11: Students logs data (OLMlets) (Chapter 5).....</b>	<b>251</b>
<b>Appendix 12: Students logs data (OLMlets) (Chapter 6).....</b>	<b>257</b>
<b>Appendix 13: Semi-structured Interview (Chapters 6 &amp; 7).....</b>	<b>259</b>
<b>Appendix 14: Semi-structured interview (Students quotes classified by themes) .....</b>	<b>261</b>
<b>Appendix 15: Students logs data (OLMlets) (Chapter 7).....</b>	<b>285</b>

# List of Figures

Figure 1-1. Tobias and Everson (2002) Framework of componential model of metacognition. ....	1
Figure 2-1. Different representations of learner models, such as a skill meter, network, radar plot, tree map, and word cloud (Bull et al., 2016b). ....	27
Figure 3-1. Examples of visual variables. ....	42
Figure 3-2. Examples of Gestalt principles. ....	43
Figure 3-3. Arrangement is used in the skill meter to indicate topics that include inconsistent data (i.e. misalignment) in an OLM. ....	49
Figure 3-4. The size of dashed line is used in a network to indicate topics that include inconsistent data (i.e. misalignment) in OLM. ....	50
Figure 3-5. Opacity is used in a network to indicate topics that include inconsistent data (i.e. misalignment) in an OLM. ....	51
Figure 3-6. The size of dashed line is used in the tree map to indicate topics that include inconsistent data (i.e. misalignment) in an OLM. ....	52
Figure 3-7. Opacity is used in the tree map to indicate topics that include inconsistent data (i.e. misalignment) in an OLM. ....	52
Figure 3-8. Blur used in a word cloud to indicate topics that include inconsistent data (i.e. misalignment) in an OLM. ....	53
Figure 3-9. The size of the dashed lines is used in a radar plot to indicate misalignment in the OLM. ....	54

Figure 3-10. Screenshot of OLMlets interface: selecting the course (top), selecting the topic (bottom). .....	57
Figure 3-11. The standard skill meter within OLMlets showing details of a misconception in the course of Java Programming. ....	58
Figure 3-12. Flow chart of the calculation of the weighted mechanism for a single question.....	60
Figure 3-13. The standard skill meter within OLMlets showing misconception details in the Java Programming course.....	61
Figure 3-14. An example of a question and a self-assessment in the extended OLMlets.....	63
Figure 3-15. Combined view.....	66
Figure 3-16. visualising alignment using separated models. ....	67
Figure 3-17. Patterns of viewing the model based on the number of users in each group. ....	69
Figure 3-18. Expandable model, where the models are combined.....	71
Figure 3-19. Expandable model where models are expanded. ....	71
Figure 3-20. Shows how the skill meter (expandable model) is affected while the student answered seven questions based on Table 3-3. ....	74
Figure 3-21. Text-based learner model within the OLMlets shows both (Computer's Beliefs, My Beliefs) specified in Study (Chapter 7). ....	75
Figure 4-1. Visualising alignment using six different methods of visual variables applied to a skill meter learner model.....	84
Figure 4-2. Scale of students' preferred method from most to least preferred.....	92

Figure 5-1. CALMsystem OLM (Kerly, Ellis and Bull, 2008) conversation fragment between the system and the learner to resolve a conflict occurred in a topic called Friction using chatbot. ....	101
Figure 5-2. Skill meter visualisation (Control condition).....	104
Figure 5-3. The combined model skill meter uses the opacity of the colour to visualise alignment.....	105
Figure 5-4. The expandable model when the model is combined. ....	106
Figure 5-5. The expandable model when the model is expanded. ....	106
Figure 5-6. Pre-questionnaire items where different views were found from the learning section before OLMlets use. ....	114
Figure 5-7. Pre-questionnaire item for confidence section where different views among students was found before OLMlets use. ....	116
Figure 5-8. Post-questionnaire item of learning section where different views .....	118
Figure 5-9. Post-questionnaire items for confidence section where different views were found after OLMlets use. ....	120
Figure 5-10. Students' views of the option to expand the models. ....	122
Figure 5-11. Box plots with error bars for students' confidence gain for the three groups of OLMlets users. ....	127
Figure 5-12. Box plots with error bars for the three groups of OLMlets users for number of questions answered. ....	130
Figure 5-13. Box plots with error bars for the three groups of OLMlets users for number of times the model was viewed. ....	131

Figure 6-1. The control condition within OLMlets specified by a participant from this study.....	141
Figure 6-2. The expandable model condition in the combined view. ....	142
Figure 6-3. The expandable model in both views (a) combined and (b) expanded specified by participant from this study.....	142
Figure 6-4. Two-way ANOVA interaction graphs between the two independent variables for bias 1 (a) and bias 2 (b). ....	152
Figure 6-5. Number of times the low-achieving control-group students visit the whole question set. ....	163
Figure 6-6. Number of times the low-achieving students from the expandable model condition visited each question set in its entirety. ....	164
Figure 6-7. Number of times the question sets were completed for the high-achieving students from the control condition.....	165
Figure 6-8. Number of visits to question sets for high-achieving students from the expandable model condition.....	166
Figure 7-1. The expandable model shown in the expanded view specified in this study by student S25. ....	176
Figure 7-2. A text-based learner model within OLMlets shows both (System's Beliefs, My Beliefs) specified in this study by student S25.....	176
Figure 7-3. Number of visits of questions for low-achieving students from the expandable model condition.....	188
Figure 7-4. Number of visits of whole questions for low-achieving students from the text-based model condition.....	189





# List of Tables

Table 3-1. Visual variables properties (Demmans Epp and Bull, 2015). .....	45
Table 3-2. Descriptions of the visual variables with examples of levels of misalignment. ....	47
Table 3-3. The system calculation based on the weighted questions of seven questions answered. ....	72
Table 4-1. Percentage of the use of different methods to represent only uncertainty according to the summary of 50 visualisations from different fields (Demmans Epp and Bull, 2015). ....	82
Table 4-2. Applying Thurstone's paired comparison method for participants' preferred method of visualising alignment. ....	90
Table 4-3. Total number of responses for each visual variable divided by total responses.....	91
Table 4-4. Normalising data using Z-transform for each visual variable.....	92
Table 4-5. Descriptive statistics for the preference for each visual variable. ....	94
Table 4-6. A 6 × 6 matrix showing the z-value of Wilcoxon signed-rank assigned to the top row for the higher mean rank for the associated visual variable.....	95
<i>Table 5-1. Number of students participating in each condition.....</i>	<i>111</i>
Table 5-2. Shows the mean rank for questionnaire items based on the difference between the three groups and the result of the Kruskal-Wallis test. The significant items are identified using bold text. ....	113

Table 5-3. The mean rank for each group on the pre-questionnaire items related to their confidence and the significant items is identified using the bold text.....	115
Table 5-4. The mean rank for each group on the post-questionnaire items related to their learning.....	117
Table 5-5. The mean rank for each group on the post-questionnaire items related to their confidence. ....	118
Table 5-6. Questionnaire items for both treatment groups about the interpretation of visualising alignment. ....	121
Table 5-7. Descriptive statistics for pre- and post-tests for both non-OLM-users and OLM-users including the sub-groups of OLM-users. ....	122
Table 5-8. Descriptive statistics for pre- and post-confidence for non-OLMlets users and OLMlets-users, including the three conditions of OLM-users. ....	124
Table 5-9. Post-hoc analysis using Bonferroni method for students' confidence gain with significant p-value indicated in bold text.....	128
Table 5-10. Descriptive statistics for students' system use for OLM-users.....	129
Table 5-11. Post-hoc analysis using Bonferroni method for number of questions answered with significant p-value indicated in bold text. ....	130
Table 5-12. Post-hoc analysis using Bonferroni method for number of times the model was viewed with significant p-value indicated in bold text.....	131
Table 6-1. Number of participants in each condition classified by students' ability level.....	140
Table 6-2. Number of questions, difficulty rating based on instructor's perception for the course Java programming. ....	146

Table 6-3. Bonferroni-corrected paired sample t-test for low-achieving students for students' bias, significant difference indicated in bold. ....	158
Table 6-4. Bonferroni-corrected Wilcoxon signed rank test for low-achieving students for students' performance, significant difference indicated in Bold. ....	158
Table 6-5. Bonferroni-corrected paired sample t-test for high-achieving students for students' bias. ....	160
Table 6-6. Bonferroni-corrected paired sample t-test for high-achieving students for students' performance. ....	160
Table 6-7. Median and inter-quartile range of number of model views for both ability levels in both conditions. ....	161
Table 6-8. Median and inter-quartile range of number of questions answered for both low-and high-achieving students for both conditions. ....	162
Table 7-1. Number of participants in each condition classified by the median score of students' performances in Set 1. ....	174
Table 7-2. Percentage of questions, difficulty rating based on instructor's opinion and the percentage of correct responses of student answers for each Bloom Taxonomy category. ....	178
Table 7-3. Students bias for low- and high- achieving students in Set 1 and Set 2 for both conditions (text-based learner model, expandable learner model). ....	182
Table 7-4. Bonferroni-corrected Wilcoxon Signed Ranks Test for students performance from Set 1 to Set 2, Significant increase from Set 1 to Set 2 is indicated as bold and italic. ....	184

Table 7-5. Mean and standard deviation for low-and high-achieving students in both conditions. .... 186

Table 7-6. Mean and standard deviation for number of times the model was viewed for both low-and high-achieving students in both conditions. .... 186

# Publications

1. **Al-Shanfari, L.**, Demmans Epp, C., Baber, C., & Nazir, M. (2018). Visualising Alignment to Support Students' Judgement of Confidence in Open Learner Models. *User Modelling and User-Adapted Interaction Journal*. (Submitted) (Chapter 6).

**Contribution of each author in the manuscript:**

Al-Shanfari, L: Conceived and planned experiments, carried out the experiment, analysed the data and wrote the manuscript.

Demmans Epp, C: provided suggestions for the structure of the manuscript, provided critical feedback and helped shape the research, analysis and the manuscript.

Baber, C: Verified the results of the analysis and provided critical feedback.

Nazir, M: Calculated the level of difficulty of the assessment and experiments held in her class.

2. **Al-Shanfari, L.**, Baber, C., & Demmans Epp, C. (2017). Student preferences for visualising uncertainty in open learner models. In E. André, R. Baker, X. Hu, M. M. T. Rodrigo, & B. du Boulay (Eds.), *Artificial Intelligence in Education. Lecture Notes in Computer Science* (Vol. 10331 LNAI, pp. 445--449). Wuhan, China: Springer. (Chapter 5)

**Contribution of each author in the manuscript:**

Al-Shanfari, L: Designed and distributed the questionnaires, analysed the data, wrote the whole paper and presented in the conference.

Baber, C: verified the design and results of the study and provided critical feedback.

Demmans Epp, C: provided critical feedback.

3. **Al-Shanfari, L.**, Demmans Epp, C., & Baber, C. (2017). Evaluating the effect of uncertainty visualisation in open learner models on students' metacognitive skills. In E. André, R. Baker, X. Hu, M. M. T. Rodrigo, & B. du Boulay (Eds.), *Artificial Intelligence in Education. Lecture Notes in Computer Science* (Vol. 10331 LNAI, pp. 15–27). Wuhan, China: Springer. (Chapter 6)

**Contribution of each author in the paper:**

Al-Shanfari, L: Conceived and planned experiments, carried out the experiment, performed the calculation, wrote the paper and presented in the conference.

Demmans Epp, C: verified the design of the experiment, suggested some methods to analyse the data, provided critical feedback and helped shape the research, analysis and the manuscript.

Baber, C: Verified the results of the analysis and provided critical feedback.

4. **Al-Shanfari, L.**, Demmans Epp, C. & Bull, S. (2016). Uncertainty in Open Learner Models: Visualising Inconsistencies in the Underlying Data, in S. Bull, B. Ginon, M. Kickmeier-Rust, J. Kay & M.D. Johnson (Eds.), Workshop on Learning Analytics for Learners, LAK16, Edinburgh, Scotland, CEUR. (Chapter 3)

**Contribution of each author in the paper:**

Al-Shanfari, L: proposed the theoretical approaches, illustrated different design approaches, wrote the whole manuscript except section 5 and presented in the conference workshop.

Demmans Epp, C: provided critical feedback, helped in developing the theoretical framework of the paper and helped shape the research paper.

Bull, S: wrote section 5 in the paper, and provided a critical feedback

5. Bull, S, **Al-Shanfari, L.** (2015). Negotiating Individual Learner Models in Contexts of Peer Assessment and Group Learning in Proceedings of Intelligent Support for Learning in Groups (AIED 2015). CEUR, International Workshop on Intelligent Support for Learning in Groups (AIED 2015), Madrid, 23 June. (Chapter 2).

**Contribution of each author in the paper:**

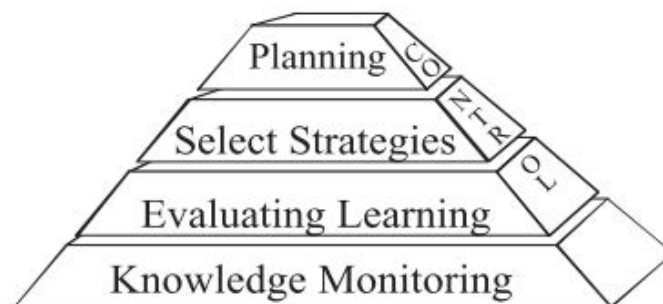
Bull, S: Wrote the whole paper.

Al-Shanfari, L: helped providing references to support the argument of the paper and presented in the conference workshop.

# Chapter 1 : INTRODUCTION

---

Independent learning requires knowledge monitoring skills (Boud, 1989) that allow students to reflect on their learning accurately. Metacognition enables students to understand and evaluate their own thinking (Flavell, 1979). Different metacognitive activities can encourage learners to reflect on their learning processes, such as self-assessment, help seeking, self-explanation, and reflection. Metacognition can help students to select, evaluate, and plan for their own learning (Flavell, 1979). Further, researchers have shown that enhancing students' awareness of their metacognitive skills can enhance students' learning (Schraw and Dennison, 1994). Tobias and Everson (2002) consider knowledge monitoring the fundamental component for other metacognitive skills. Accordingly, they proposed a theoretical framework of metacognition, as shown in Figure 1-1.



*Figure 1-1. Tobias and Everson (2002) Framework of componential model of metacognition.*

Tobias and Everson (2002) define knowledge monitoring as the ability to distinguish between what one does or does not know. The aim of the study in this thesis is to support students' knowledge monitoring skills by enabling students to observe the alignment between their confidence in the correctness of answers, based on their own self-confidence rating without external influences, (referred to in this thesis as confidence judgement) with students' actual performance, based on computer evaluation using an adaptive educational system. Misalignment arises when the computer evaluation and student confidence conflict. For example, a student could generate a correct answer while they were feeling unconfident. Further, a student could be confident that they give a correct answer, but then provide an incorrect response. Tobias and Everson explained that students who lack knowledge monitoring skills spend more time learning concepts that they have already mastered than on focusing their attention on other concepts yet to be mastered (Tobias and Everson, 2002).

Knowledge monitoring has gained attention from different researchers, interested in how to improve students' knowledge monitoring skills. From these studies, a number of theories have been proposed. One such theory suggests that giving students the ability to practice their knowledge monitoring skills can improve their monitoring accuracy, by shifting confidence judgement from reliance on their own beliefs about their abilities to experience-based judgement (Koriat, 1997). Students providing answers to questions may perform correctly, but the correct answer may represent either solid knowledge or a random guess. Students with good knowledge monitoring skills can accurately align their own confidence in the



correctness of their answers to their actual performance. Another theory related to how student behaviour is affected when assessing their knowledge monitoring skills states the following:

High certitude correct responses should remain high because the learner is certain of the response, and it is the correct one. High certitude errors represent an interesting problem since they are the point of maximum discrepancy where the learner works hardest to correct them (Kulhavy and Stock, 1989, p.294).

This theory suggests that when misalignment is detected between students' confidence and their performance, the student may seek to resolve the misalignment and this may affect their behaviour monitoring their knowledge. Researchers also argue that meta-cognitively skilled students have the capability to plan, monitor the success and the weakness of their own activities, and alter behaviour in accordance with their monitoring. Low-achieving students are less proficient at monitoring their knowledge (Royer, Cisero and Carle, 1993). Experimental studies showed that low-achieving students lack knowledge monitoring skills. Further, they tend to over-estimate their performance, which can lead them to discontinue their efforts in their learning (Koriat, 1997; Kelemen, 2000). Students' lack of knowledge monitoring skills while learning is attributed to motivation (Zimmerman, 1990). Therefore, it is important to motivate students to practice their knowledge monitoring skills, and to provide them with information that bonds their performance with their confidence.

Developing students' metacognitive skills using Intelligent Tutoring Systems (ITSs) is fruitful. Recently, researchers have developed meta-cognitively intelligent tutoring systems, 'systems that care', to support students to regulate their learning more effectively (du Boulay et al., 2010). These ITSs include monitoring functionality to enable the adaptation of learning materials and experience to students' individual needs. The ITSs rely on a specific type of user model (often called a learner model) to drive this adaptation. For example, an intelligent tutoring system could keep records of past actions in the learner model to allow students to learn from their previous experience when repeated patterns of cognitive and metacognitive occur (Hull and du Boulay, 2009). The learner model is typically hidden, which means that students cannot access their own learner model. Researchers have argued that the system should allow learners to access the learner model (Self, 1990), the main reasons being to promote metacognitive behaviour, support students with self-regulated learning, and promote learner reflection (Bull and Kay, 2013).

The aim of this thesis is to help students monitor their knowledge in terms of subject understanding and improve their knowledge monitoring skills by using Open Learner Models (OLMs). In this thesis, the research questions address the potential of incorporating students' confidence in correctness of their answers and the system evaluation of students' performance into an OLM. The approach used allows students to monitor their own learning by visualising the alignment in the context of OLMs. In this thesis, the effect of the alignment/misalignment in assessing confidence in the correctness of their answers on students' behaviour

and the effect of practicing the confidence judgement are investigated within OLMs. Designing an OLM and presenting the information in a way that could influence the enhancement of students' metacognitive skills is an essential aspect that could contribute to the field of ITSs and OLMs.

### **1.1 Research Questions:**

1. How does visualising alignment support students' confidence judgement?
2. What is the preferred method to visualise alignment for OLM users?
3. How is the information's presentation format important to the alignment information in an OLM?

### **1.2 Metacognition**

Metacognition is the ability to reflect on ones' own knowledge in that it is 'thinking about thinking' (Flavell, 1979). Metacognition can be divided into three components: metacognitive knowledge, knowledge monitoring, and the control of these processes (Pintrich, Wolters and Baxter, 2000). Flavell describes metacognitive knowledge as the whole knowledge of the individual. He continues by stating that this knowledge can be influenced by factors such as the person, the tasks the person has taken, and the strategies that the person will decide to follow to accomplish these tasks (Flavell, 1979).

Metacognitive monitoring includes metacognitive judgement, which is the ability to judge the correctness of the answers given. The third component is how the person can control activities to achieve a desired goal. Supporting students'

metacognitive judgement is the focus of this thesis. The next section provides methods to measure metacognitive judgement.

### **1.3 Measures of Metacognitive Judgement**

Monitoring students' knowledge allows them to reflect and become aware of their own knowledge. There have been a variety of methods to measure metacognitive judgement, which include the following: the self-reporting of the person's own judgement, thinking aloud (where students verbalise their thoughts while doing a task), and error detection (Pintrich, Wolters and Baxter, 2000). Error detection is the use of a text that contains some errors that the student should be able to identify, to show their ability to monitor their knowledge.

In this thesis, the method of self-reporting is used. There has been intensive research to evaluate students' knowledge monitoring ability using different assessment techniques based on students' self-reporting. Tobias and Everson (2009) provided a technique that measures the students' ability to differentiate between what they do and do not know, called the knowledge monitoring assessment (KMA). In assessment, students were asked to evaluate their ability to solve a problem with a yes or no answer before actually solving it. A comparison between their judgement and their actual performance was evaluated, and provided the following outcomes (Tobias and Everson, 2009):

- Students who stated that they knew the answer and performed correctly:  
(a) [+ +].

- Students who stated that they did not know the answer but had given a correct answer: (b) [- +].
- Students who stated that they knew the answer but performed incorrectly: (c) [+ -].
- Students who stated that they did not know the answer and indeed performed incorrectly: (d) [- -].

The knowledge monitoring score is obtained by using the simple equation [1], (Tobias and Everson, 2009)

$$KMA = \frac{(a + d) - (b + c)}{(a + b + c + d)} \quad [1]$$

The KMA score ranges between +1 and -1. The score -1 indicates that the student has poor knowledge monitoring skills and +1 shows accurate judgement.

In this thesis, the system has been extended to allow students to evaluate their confidence of the answer given, after responding to each question of the system-automated assessment. Thus, the KMA measure is not an appropriate measure to use in the context of this thesis because the confidence judgement is assessed after solving the question and there are more than two options, whereas KMA assesses students' confidence before using the system and with two confidence options.

Schraw (2009) described five measures for assessing one's own knowledge when the individual evaluates their confidence after the completion of a test item. The measurement applied in the confidence rating uses an ordinal scale from 0

(no confidence) to 100 (totally confident). The first method to measure students' confidence judgement is absolute accuracy, which is the difference between a student's confidence rating of their abilities to accomplish the task and their actual performance on that task. The higher the absolute accuracy the worse the student's knowledge monitoring skills. The second measure is relative accuracy, which is the relationship between multiple confidence judgements and performance scores. This relationship focuses on the difference of a single confidence judgement and mean confidence and the difference between performance on a single item and mean performance. Relative accuracy is usually measured using a gamma coefficient (Maki et al., 2005), used for measuring the strength of association between variables, which (in this instance) would indicate whether items that receive higher judgement are also the items that produce higher performance and whether items that receive lower judgement also produce lower performance. However, relative accuracy does not show whether the student is over- or under-confident. The other three measures that can be used are bias, scatter, and discrimination. Bias of student's confidence judgement shows the direction of the individual's ability to estimate their performance accurately. Bias was calculated as a unitised difference of the confidence judgement and the performance, using Equation (2), (Schraw, 2009), as follows:

$$\text{Bias} = \frac{\sum_1^n (\text{confidence} - \text{performance})}{n} \quad [2]$$

The bias score is always between -1 and +1. The sign magnitude for the bias score is negative for under-confidence, positive for over-confidence, and a score

of 0 indicates that students' performance was aligned with their confidence judgement. Based on Equation 2, a bias score of +1 is obtained when the student fails to answer their knowledge assessment correctly but they are confident that the answer is correct. Conversely, a bias score of -1 is obtained when the student never fails in their knowledge assessment but they are unconfident that the answer is correct for the whole assessment. The score of 0 (null score) is obtained when the student aligns their confidence to their performance.

The fourth measure is the scatter, which shows the degree of individual variability in correct and incorrect responses. The fifth measure, which is discrimination, indicates the ability of the individual to differentiate between known and unknown items. A positive sign of discrimination indicates the individual is meta-cognitively aware that their answers are correct, whilst a negative sign represents lack of awareness (Schraw, 2009).

In the context of this thesis, the bias measure is used as a proxy to represent the accuracy (or alignment) of a student's self-assessed confidence, which is called confidence judgement, to the system's evaluation of students' performance. Misalignment occurs when there is a disagreement between data sources. That is, when the system shows that the student understands the task but the student is not confident of having answered correctly, or when the system evaluation shows that the student lacks the necessary knowledge but the student is confident in his or her knowledge of the material. When a misalignment is detected, the system becomes uncertain about the state of the learner model. This uncertainty in

learner models can be represented through OLMs (Demmans Epp and Bull, 2015).

#### **1.4 Can Students Accurately Assess their own Knowledge?**

Metacognitive monitoring is a critical aspect of selecting a strategy and planning for one's learning. It provides the basis for deciding what and how long to study certain concepts. Researchers argue that metacognitive monitoring affects student self-regulation study, and thus it can affect overall learning (Thiede, Anderson and Therriault, 2003). Based on the framework of Tobias and Everson (2009) shown in Figure 1-1, knowledge monitoring skills are a pre-requisite for metacognition. Although students' ability to monitor their knowledge varies from one student to the other, several studies have confirmed that low-achieving students tend to be over-confident in their abilities, whereas high-achieving students are more accurate in their confidence (DiFrancesca et al., 2016; Hacker et al., 2000; Nietfeld, Cao and Osborne, 2005). This suggests that high-achieving students are more consistent and that low-achieving students face many challenges and need more differentiated support. Some researchers suggest that low-achieving students are unaware of, and un-skilled in, metacognitive skills (Ehrlinger et al., 2008). However, other researchers suggest that while low-achieving students are unskilled, they may be aware that they lack metacognitive skills (Miller and Geraci, 2011).

To provide assistance for low-achieving students to improve their knowledge monitoring skills, it is necessary to know the factors that might influence knowledge monitoring skills. For example, test difficulty affects students'



confidence judgement, with students tending to be over-confident with more difficult tests (Nietfeld, Cao and Osborne, 2005; Schraw and Dennison, 1994). Domain knowledge is also an important factor. Students who are familiar with the domain become more confident and can assess their knowledge accurately (Nietfeld and Schraw, 2002).

Providing students with feedback is important to enable them to reflect on their awareness about their level of knowledge on an assigned task and their knowledge monitoring skills. For example, according to Callender, Franco-Watkins and Roberts (2016), providing feedback at the global level (i.e. for the whole test) about performance and judgement of confidence led to improved students' performance and accuracy in their confidence judgement in a second test, in comparison to students who did not receive feedback after their first exam. Further, an improvement was identified in students' performance and confidence judgement when feedback was provided at the local level, on a question-by-question basis (Renner and Renner, 2001). Training students to monitor their knowledge through feedback led to improvements in knowledge monitoring (Huff and Nietfeld, 2009). However, practising knowledge monitoring skills without receiving feedback was not enough to result in improvement in monitoring their knowledge (Nietfeld, Cao and Osborne, 2005).

## **1.5 Study Objectives and Approach**

Facilitating students' knowledge monitoring skills in an OLM is fundamental for supporting other metacognitive skills for more active learning. The main aim of this thesis is to investigate how students' confidence in what they know (learned)

and what they do not know (unlearned) can be supported by visualising model alignment.

To fulfil the aim of this thesis, I have designed and built a model alignment and its representation in the base system (i.e. OLMlets OLM) to extend its features. OLMlets is programmed using PHP scripting language, which is connected to the MySQL database. OLMlets is hosted on the Apache web server and accessed through the link (<http://olmlets.bham.ac.uk>). Visualise alignment is represented between the computer's evaluation of the students' performance in the system's automated assessment and students' confidence that the answers given are correct. The research objectives are as follows:

- To explore students' preferred method of visual variable to visualise alignment (Chapter 4).
- To investigate the impact of the alignment based on pre- and post-confidence, following an OLM use (Chapter 5).
- To investigate the impact of the alignment while students practice the confidence judgement within the OLM (Chapter 6).
- To investigate the importance of the information presentation of the alignment in the OLM (Chapter 7).
- To explore students' behaviour when viewing model alignment in an OLM (Chapters 5-7).

## **1.6 Chronology of the work**

The thesis presents four studies to answer the research questions of this thesis. All studies were conducted after receiving the ethical approval from the University of Birmingham Ethics Committee (reference ERN\_15-0684P). The ethical approval involved submitting a full description of the studies together with the participation information sheets, to the university Ethics Board, who then issued a reference number to indicate that the proposal was ethically sound.

In this thesis, the evaluation studies were conducted using real class settings of undergraduate students from the School of Electronic, Electrical and Systems Engineering at the University of Birmingham (United Kingdom) and the Sultan Qaboos University (Oman). Courses related to computer programming and mathematics were integrated in the OLMlets system. Students used the OLMlets system as an independent system to help with their exam preparation.

The thesis was based on the research of Demmans Epp & Bull (2015), about uncertainty representation in visualisation in the context of OLMs. Different visual variables were used to indicate uncertainty in the data using different forms of OLM visualisation. The theory of the thesis was based on Kulhavy and Stock (1989), whereby the greater the discrepancy between student confidence about the correctness of the answer and their response, the more motivated the student is to reveal this discrepancy and more time is spent processing the feedback. Prior OLM research (Bull and Pain, 1995) investigated the impact of allowing students to align their confidence to their performance using a negotiated OLM, which uses a form of discussion between the student and the computer until

agreement is reached to maintain the accuracy of the model. From the research discussed above, I was interested to implement the proposed method of Demmans Epp and Bull, (2015), to visualise alignment in a simple OLM to investigate the impact of the alignment on students' confidence judgement and their behaviour when using the system.

I have conducted a pilot study to inform the design of the representation of visualising alignment in OLMs, which is discussed in Chapter 3, Section 3.3.5.

I decided to implement one form of visual variable (suggested by Demmans Epp and Bull, 2015)—which is the opacity of the colour to indicate the alignment in OLMs—using a control condition study. This was the first study presented in Chapter 5.

Demmans Epp and Bull (2015) used different visual variables, conducting an analysis of 50 visual variables to represent uncertainty in the data effectively. I was interested to explore the preferred method using six from those visual variables, which were added marks, arrangement, opacity, numerosity, orientation, and size. This was done to help system designers to inform the design of the system when students preferred method is important. This was the second study presented in Chapter 4.

The third study was to investigate how visualising alignment impacts students' bias of confidence judgement using students' preferred method using a control condition study (Chapter 6). Surprisingly, students preferred the opacity of the

colour as their preferred method to visualise alignment in OLM, which was implemented in Study 1 (Chapter 5).

Although prior OLMs use text-based representation widely in their design to include student's self-assessment, there was an argument that presentation methods such as visual or textual may influence students' metacognitive judgement (Burkett and Azevedo, 2012). Thus, Study 4 was conducted to investigate the impact of presentation format on the alignment information using two representations: graphical-based, using students' preferred method explored in Study 2; and text-based, which was used in prior OLMs, such as in Mr. Collins (Bull and Pain, 1995).

## **1.7 Contributions**

This thesis extends prior research that uses OLMs to promote learner reflection, by giving students the ability to compare the confidence judgement between two sets of beliefs (i.e. alignment). First, students' judgement about the extent that students' met the teacher's criteria for answering problems are aligned with the teacher's criteria defined for each task, for example Reflect OLM (Kay, Li and Fekete, 2007). Second, student self-assessment of their confidence that the answers are correct is aligned with the computer evaluation of students' performance on the task for example, Mr. Collins, CALMsystem, and VisMod (Bull and Pain, 1995; Kerly and Bull, 2008; Zapata-Rivera and Greer, 2004).

As there is no prior research on the impact of visualising alignment on students' knowledge monitoring skills and confidence using a control condition in the

context of OLMs, this thesis provides a practical implementation of real class studies that investigate the impact of providing students with information about themselves in their interaction with the system. More specifically, the research seeks congruence between the computer's view of the learner and the learner's perception of themselves. The findings contribute to the literature in the fields of Intelligent Adaptive Systems and OLMs, as they demonstrate the importance of visualising alignment in OLMs to support students' knowledge monitoring skills, especially for low-achieving students who need more differentiated support (discussed further in Chapters 6 and 7).

The second contribution of this thesis is to the fields of intelligent adaptive systems and OLMs, in that it illustrates a variety of design approaches using different methods of visual variables to visualise alignment that can be applied to different representations of the learner model. This thesis explores students' preferences for styles of visualising alignment applied to one simple representation (skill meter). To date, indicating students' preference style to visualise alignment is a novel act in OLMs. These design approaches can develop guidelines that can inform designers of OLMs and ITSs when visualising alignment is important for facilitating students' knowledge monitoring skills (Discussed in Chapter 4).

The third contribution of this thesis is that information presentation of the underlying learner model is an important aspect to consider when visualising alignment to facilitate students' knowledge monitoring skills.

Before explaining the approaches of the OLM to support students' metacognitive skills, the next section provides details of the theories related to metacognition and, more specifically, to students' knowledge monitoring from the field of education and psychology.

## **1.8 Thesis Summary**

This thesis addresses three research questions related to the presentation of alignment information in open learner models and the impact of adding this information on students' metacognitive skills. The following sections summarise the research questions and how these are approached in this thesis.

1. How does visualising alignment support students' confidence judgement?

It is argued that OLMs promote students' metacognitive skills (Bull and Kay, 2013). Some OLMs allow students to reflect on their own answers and facilitate metacognitive skills, such as help-seeking (Roll et al., 2011) and problem solving (Mitrovic and Martin, 2007). Researchers argue that supporting students' confidence judgement may improve metacognitive skills to aid in selecting the appropriate problem to solve (Mitrovic and Martin, 2007; Roll et al., 2011). This reflects the importance of knowledge monitoring skills in promoting metacognitive skills (Tobias and Everson, 2009).

Although some OLMs allow students to align their confidence judgement compared with the computer or the teacher's viewpoint (Bull and Pain, 1995; Kay, Li and Fekete, 2007; Kerly and Bull, 2008; Zapata-Rivera and Greer, 2004), these

systems used either the negotiation method (i.e. both the computer and the student should agree to resolve the misalignment based on some form of discussion) (Bull and Pain, 1995; Kerly, Ellis and Bull, 2008), or the simple OLM Reflect-OLM (Kay, Li and Fekete, 2007). These systems do not use a control condition to investigate the impact of the alignment on students' knowledge monitoring skills.

Moreover, researchers argue 'simply presenting an inspectable OLM by itself may not be an effective way to support self-assessment and additional scaffolding may be necessary' (Long and Alevan, 2013c).

Therefore, it is still an open question whether OLM can help students to be more accurate in assessing their own knowledge. Hence, this thesis provides an investigation of this open research question and provides a form of visualisation of the alignment, demonstrating that the student can easily compare the two sets of beliefs of the computer evaluation of students' answers and students' confidence about the correctness in the answers given. This can help designers of OLMs and ITSs to consider visualising alignment when their goal is to support students' confidence judgement.

## 2. What is the preferred method to visualise alignment for OLM users?

Exploring students' preferred style to visualise alignment in OLMs is novel. Although some studies outside the educational context show that participants may prefer a method that is not necessarily the most effective (Boukhelifa et al., 2012), there are researchers who believe that providing users with their preferred



method can help them understand the information depicted by those visualisations (Gerharz and Pebesma, 2009). Knowing students' preferences in visualising alignment could provide guidelines to designers about which approaches students may find beneficial.

### 3. How is information presentation format important in alignment information in OLMs?

Prior research argues that the presentation method (i.e., visual or textual) may influence students' metacognitive judgement (Burkett and Azevedo, 2012). Based on the first question about how visualising alignment can support students' metacognitive skills, this thesis provides an investigation into whether the representation of the alignment is important to facilitate students' knowledge monitoring skills when it is represented using a textual or visual description. Representing alignment information using text-based description has been used in prior OLMs such as Mr.Collins and ReflectOLM (Bull and Pain, 1995; Kay, Li and Fekete, 2007). In Mr.Collins (Bull and Pain, 1995), students were willing to negotiate with the system to change the learner model when they did not agree with the content to provide a more accurate learner model. The Self-assessment Tutor (Roll et al., 2011) used text, allowing students to assess their confidence before and after solving the problem. The Self-assessment Tutor did not solve the problem of students' over-confidence. (These studies used qualitative analyses only (Bull and Pain, 1995; Kay, Li and Fekete, 2007) or self-assessment combined with other metacognitive skills (Roll et al., 2011; Long and Aleven, 2013c)). Based on prior research, this thesis provides an investigation on the impact of visualising

alignment in OLMs and how students' confidence judgement will differ according to the mode of the alignment representation (skill meters vs. text) explored in (Chapter 7). This can provide OLM designers and other interactive learning environments with the most appropriate representation to use for facilitating students' knowledge monitoring skills.

## **1.9 Outline of Chapters**

This section provides an overview of the structure of the rest of this thesis.

Chapter 2 — This chapter provides the background and related work for this thesis through a review of the framework of the OLM approach and how OLMs supports students' metacognitive skills. More precisely, it provides examples of OLMs that have implemented knowledge monitoring skills in their systems together with the type of OLM representation. The chapter also describes the related work of OLMs that used the alignment between the students' confidence and the computer's view about the student performance, highlighting the current gap within the field.

Chapter 3 — This chapter focuses on information visualisation and how visual variables can play an important role in indicating the misalignment to provide the user with an overview of the inconsistency in the learner model without focusing their attention on the data. Examples of how indicating inconsistency (e.g. misalignment) in the learner model using different visual variables are applied to different visualisation of the learner models. Finally, the system design and architecture of the existing system is described (i.e. the OLMlets system). The learner

modelling process of the base system of OLMlets OLM is explained, together with how I modelled and built the model alignment and its representation of the learner model in the extended version of OLMlets.

Chapter 4 — This chapter presents a qualitative study that was performed to explore the preferred method to visualise alignment in OLMs using students' self-reports. In this study, students were asked to choose one of the visual variables among six other variables applied in simple learner model visualisation (e.g. skill meter). By doing this, the students' preferred method of indicating alignment using visual variables in a simple learner model visualisation (skill meter) was explored.

Chapter 5 — In this chapter, an empirical study was performed to investigate the impact of visualising alignment on students' confidence gain as well as their interaction with the system. In this study, students were asked to complete a pre-questionnaire and to solve a pre-test before they used the system. At the end of the study, students were asked to complete a post-questionnaire and to solve a post-test. By doing this, the impact of adding alignment in the visualisation of the learner model was explored.

Chapter 6 — In this chapter, another empirical study was conducted to investigate the impact of visualising alignment on students' confidence judgement using a control condition. In this study, the impact of visualising alignment on students' confidence judgement and action was explored.

Chapter 7— In this chapter, a follow-up study was conducted to investigate how the representation of alignment is important. Two versions of the alignment were developed: text-based and visual-based. This study used the same students from the study discussed in Chapter 6.

Chapter 8 — This chapter summarises the findings of this thesis and discusses its limitations, potential contributions, and future work.

# Chapter 2 : Open Learner Models

---

This chapter presents an overview of Intelligent Tutoring Systems (ITSs) and Open Learner Models (OLMs). It provides examples of OLMs that use components of metacognition to support student learning. It also illustrates different OLMs that use the alignment between students' confidence and the computer's or teacher's own beliefs about the student's performance using a text-based or visual learner model.

## 2.1 Intelligent Tutoring Systems

The monitoring of student knowledge can be performed by software, teachers, or students. Intelligent Tutoring Systems are a type of software that monitors student knowledge. They include monitoring functionality to enable the adaptation of learning materials and experiences to students' individual needs (Woolf, 2009).

Intelligent Tutoring Systems rely on a specific type of user model, often called a learner model. The architecture of ITSs involves four main mechanisms: domain model, learner model, tutoring model, and the user interface (Nwana, 1990).

The source of knowledge that evaluates student performance and diagnoses errors is known as the domain model, or the expert model. As the student interacts with the system, all inference and system beliefs about the student are stored in the learner model (Woolf, 2009). The learner model contains specific information that the system has deduced about the learner. The learner model may comprise

information such as the level of knowledge of the learner, possible misconceptions, emotions, and competencies. Information in these models guides adaptivity in ITSs (Holt et al., 1994).

Most tutoring systems contain an outer loop that is responsible for selecting the next learning task and for executing each complex task as a whole (VanLehn, 2006). The minimal requirement for the tutoring system to be considered intelligent is the identification of an inner loop that provides adaptive support as the student is working on a task, and follows each step the student has taken in solving the task. The inner loop is responsible for providing feedback, assessing the students' evolving knowledge, and updating the learner model each time the student interacts with the system (VanLehn, 2006). Accessing the learner model (making it inspectable) could promote students' metacognitive skills, by allowing the learner to think about their own learning, thus promoting reflection (Bull and Kay, 2013). The next section provides an overview of the approach used in this thesis.

## **2.2 Open Learner Models**

Open Learner Models are underlying learner models that are externalised to the learner through visualisation (Bull and Kay, 2016). In Open Learner Models, students gain responsibility for their learning by accessing the learner models. Opening the learner model to students can have benefits for students' learning and their metacognitive skills, when compared with a closed learner model (Bull and Kay, 2013). For example, OLMs could promote learner reflection by externalising the learner model to students. By providing open access to students, this would

provide the capability of facilitating planning and monitoring of learner knowledge, facilitating collaboration among learners and improving its accuracy, by allowing the learner to contribute to his or her own learner model (Bull and Kay, 2016). Moreover, allowing the learner to access the learner model may increase trust in the content of their own learner model (Ahmad and Bull, 2008).

There are different types of OLM, each with their own unique features. Some allow the learner to only inspect the learner model (Bull and McKay, 2004; Bull, Jackson and Lancaster, 2010), whilst others give permission to edit the content of the learner model directly (Kay, 1997; Mabbott and Bull, 2006) or to negotiate the system to maintain learner model content, where both the system and the learner have control of the content of the learner model (Bull and Pain, 1995; Bull and Al-Shanfari, 2015; Dimitrova, 2003; Kerly, Ellis and Bull, 2008).

The learner model can be constructed using different learner modelling approaches, such as the following: Bayesian networks—for example VisMod (Zapata-Rivera & Greer, 2004); constraint-based modelling—for example SQL-Tutor (Mitrovic & Martin, 2007); and simpler weighted algorithms—for example OLMlets (Bull, Jackson & Lancaster, 2010).

Different OLM approaches use a wide variety of visualisation methods. These range from simple ones such as skill meters (Bull et al., 2016a, 2016b; Bull and Mabbott, 2006; Corbett and Bhatnagar, 1997; Duan, Mitrovic and Churcher, 2010; Mitrovic and Martin, 2007) to more complex methods, such as networks, radar plot, word cloud, (Bull et al., 2016a, 2016b), or hierarchal tree structures

(Duan, Mitrovic and Churcher, 2010; Kay, 1997; Mabbott and Bull, 2004). Figure 2-1 shows examples of these types of OLM. The most common and simple visualisation is a skill meter (top left of Figure 2-1). In skill meters, the student's level of knowledge is indicated by the filling of the bar, and it can be useful with a minimal number of topics. However, one disadvantage of the skill meter is that when the domain includes a larger number of topics, the learner needs to scroll to view the topics. Network visualisation, top right of Figure 2-1, displays more topics; however, there might be difficulty in reading when nodes are near each other. The level of knowledge in the network is indicated with different variables (such as size and colour) to indicate the level of knowledge of the learner. When the node is bright and large, this indicates that the learner has achieved higher level of knowledge.



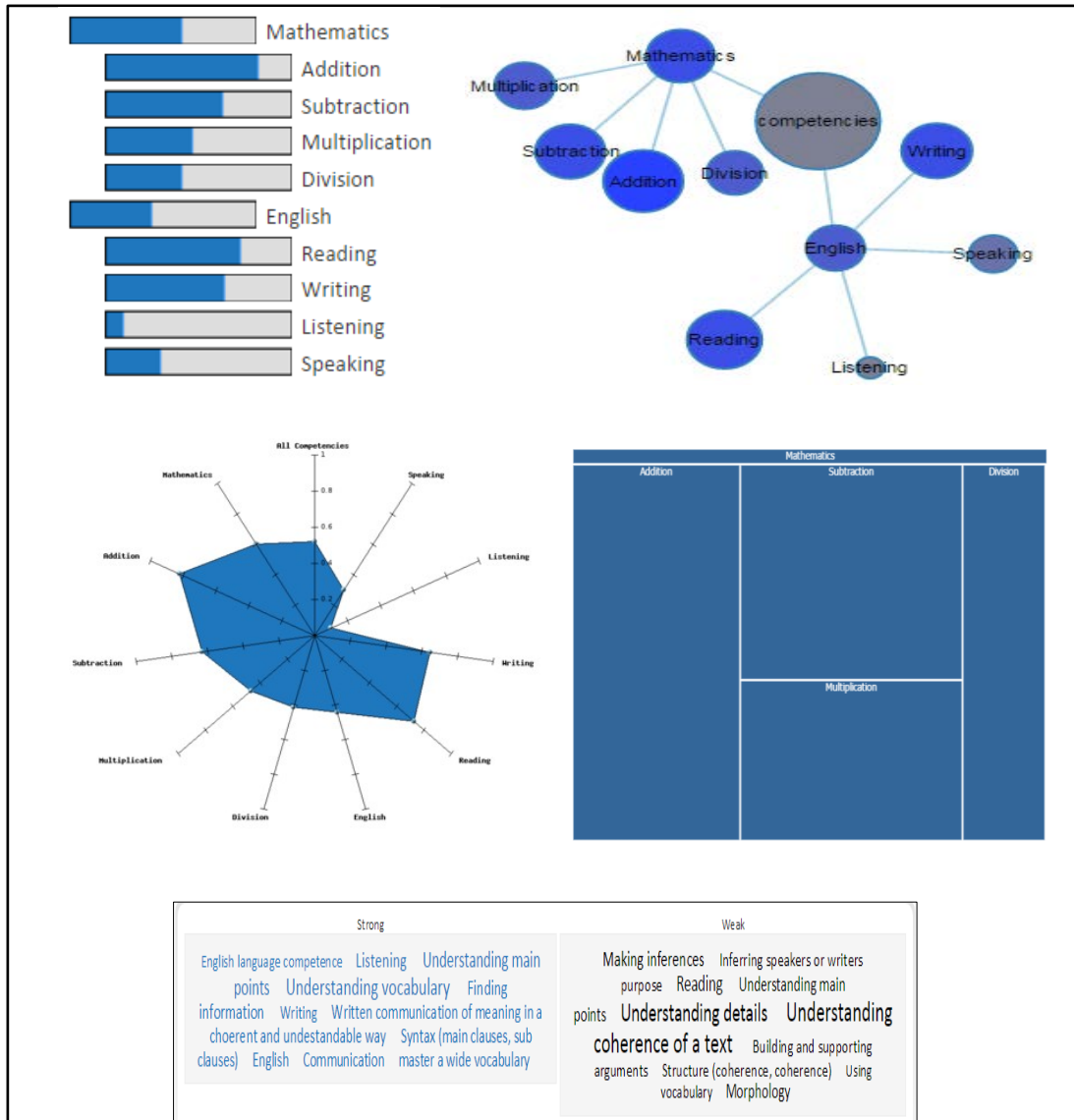


Figure 2-1. Different representations of learner models, such as a skill meter, network, radar plot, tree map, and word cloud (Bull et al., 2016b).

The radar plot, bottom left in Figure 2-1, can also visualise the learner level of knowledge and can include a greater number of topics in a smaller space than the skill meters. However, learners do not see the overall domain structure on one screen.

When the domain contains a large number of topics, the tree map could be considered, as it allows the learner to view the topics and sub-topics. In the tree map, the learner could access the sub-topics by clicking on the parent topic. However, whilst the learner could take advantage of exploring a wide number of topics, they will not be able to compare topics from different parts of the tree map. The size of the rectangle of each topic indicates the level of understanding of the learner. The larger the size of the rectangle, the higher the level of knowledge. The word cloud at the bottom of Figure 2-1 separates the topics into two boxes: in the strong box, the larger the size of the word indicates stronger skills, whereas the larger size in the weaker box are weaker skills. However, topics that are at the borderline of both strong and weak topics could be difficult to read, because of their small size.

Allowing the learner to access the learner model through visualisation could lead them to reflect on different aspects of their learning and enhance their metacognitive skills.

## **2.3 Supporting Students' Metacognitive Skills in Open Learner**

### **Models**

Metacognition can support students in their learning (Schraw and Dennison, 1994). Researchers argue that the major challenges for learners lie in monitoring and regulating their cognitive and metacognitive processes during learning (Azevedo and Aleven, 2013). Based on Winne and Hadwin's model of self-regulated learning (SRL), Greene and Azevedo (2007) suggest that learning occurs

in four basic phases: task definition, goal setting and planning, studying tactics, and adaptation to metacognition. Intelligent tutoring systems provide adaptive support for individual students and highlight the gaps to the learner, fostering their metacognitive and self-regulated behaviour (Azevedo and Alevan, 2013).

Different metacognitive skills are supported using ITSs, such as reflection (Kay, Li and Fekete, 2007), help seeking (Roll et al., 2011), self-awareness (Bull, Jackson and Lancaster, 2010), and self-assessment (Mitrovic and Martin, 2007). Each of these metacognitive activities involves the processes of thinking, monitoring, and being aware about one's own learning (Flavell, 1979). For example, in SQL-Tutor (Mitrovic and Martin, 2007), there is a constraint-based tutor that supports undergraduate students to learn the dominant database query language. As discussed in Section 1.4, students' metacognitive knowledge can be influenced by different factors, namely person, task, and strategy (Flavell, 1979). This will be explained here, based on the mechanism of the ITS and by applying the Winne and Hadwin's model of SRL (Greene and Azevedo 2007), where learning occurs in four phases. In SQL-Tutor (Mitrovic and Martin, 2007), phase one of task definition in the SRL model was applied by allowing the student (person variable) to select and assess the difficulty of the task. Phase 2 was applied when the student set the goal and planned how to accomplish the task, by either completing the problem (task variable) or have the ability to abandon the problem. The student can select another problem (strategy variable), applying phase three of the SRL model. The last phase of the model SRL (adaptation to metacognition), was applied where SQL-Tutor asks the student to clarify why they abandoned a problem.

This is achieved by selecting one of three possible options: 'the problem is too easy', 'the problem is too difficult', or 'I need to work on different type of problem'. Students differ in the amount of knowledge that they acquire about tangible resources for accomplishing the task. Students with good knowledge monitoring skills would select the appropriate problems that may help them understand the subject. The SQL-Tutor uses a skill meter with a green fill-colour to represent the learned concept, red for still learning, and white for concepts that are not yet covered. The learner model in SQL-Tutor does not show students information about the accuracy of their metacognitive judgement. In a comprehensive evaluation, it was evident that novice students benefited from using SQL-Tutor, which improved their learning and triggered a more positive attitude in high-achieving students (Hartley and Mitrovic, 2002). These benefits to the student could be because the SQL-Tutor allows the student to assess the level of difficulty of the problem, promotes their reflection, and provides them with an awareness about their strengths and weaknesses.

Knowledge monitoring skills are a pre-requisite of metacognition (Tobias and Everson, 2009). Monitoring accuracy could affect students' long-term retention in learning where over-confidence could lead to poor study choices that subsequently affect learning (Dunlosky and Rawson, 2011). Thus, it is significant for ITSs to support the acquisition of metacognitive skills, such as knowledge monitoring skills.

Different ITSs use different methods to measure student knowledge monitoring skills, as explained in Section 1.5. For example, the 'think aloud' method was

used in a Fraction Addition Tutor where students were required to think aloud to determine whether they correctly interpreted the learner model (Stampfer et al., 2011). The Intelligent Novice Tutor (Mathan and Koedinger, 2005) uses the error-detection method to identify errors from student answers. Subsequently, the system highlights the errors and provides an interface for error-correction. The student can then compare the two interfaces to identify discrepancies between the error and the error-correction outcome, which can enhance their metacognitive skills. Self-report is widely used to assess student confidence in their answers to the system's automated assessment (Bull and Pain, 1995; Kerly, Ellis and Bull, 2008; Long and Alevan, 2013b).

Next, some examples of tutoring systems that use self-assessment and other metacognitive elements to help students monitor their knowledge and reflect upon that on their learning are briefly described.

### **MetaTutor**

MetaTutor (Azevedo et al., 2009) is hyper-media learning environment, designed to scaffold students' metacognitive, self-regulated learning process. MetaTutor uses an animated pedagogical agent to prompt and provide feedback to students' use of different metacognitive skills, such as planning, metacognitive monitoring, and learning strategies, when learning complex science topics. Including different components of metacognition in MetaTutor has been shown to enhance students' awareness and help students to monitor their progress to achieve their goals, enhancing their conceptual comprehension (Azevedo et al., 2011). It is also

shown to support learning strategies for low-performing students by using note-taking (Trevors et al., 2014). MetaTutor allowed students to set their goals and monitor and control their own learning, which positively influences student-learning achievement (Duffy and Azevedo, 2015).

### **The Reflection assistance**

The MIRA learning system (Gamma, 2004) focuses on different metacognitive skills such as problem selection, knowledge monitoring, planning, and evaluation. The Reflective Assistant (RA) in the MIRA system allows students to think about themselves as problem solvers and reflect on their learning, through reflective activities to support students to succeed in solving the problem. The MIRA system builds a metacognitive profile of the student, showing the accuracy of the student's knowledge monitoring accuracy and the bias of student's confidence judgement in relation to their performance. Including the RA model in the MIRA system demonstrates a positive effect on student's metacognitive skills (Gamma, 2004).

### **Self-assessment Tutor**

Self-Assessment Tutor (Roll et al., 2011) supports students' help-seeking behaviour. Students judge their ability to solve the problem using a text-based format from a drop-down list, by answering the question 'can you solve this problem without making an error?' with 'yes', or 'no, I need a hint'. Students update their self-assessment after solving the problem by answering the question 'did you correctly evaluate your knowledge?' and choosing from either of the two

responses ‘no, I thought I knew it but was wrong’ or ‘no, I knew more than I predicted’. The study used to evaluate the Self-assessment Tutor indicated that students improved their ability to judge their solutions to questions, that they solved them correctly, but did not lose their overconfidence. The study also suggested that students’ help-seeking behaviour improved, which may have been due to the effect of the metacognitive practice of the self-assessment. However, this shows that embedding the self-assessment in the ITSs may enhance other metacognitive skills such as help seeking, this study did not provide evidence to support that aligning students’ confidence judgement before and after solving the problem could improve their accuracy in their self-assessment skills on both their strengths and weaknesses. Therefore, more studies are required to evaluate the effect of students’ confidence judgement practice in the ITSs environment and their transferred effect on metacognitive skills.

The next section presents some examples of OLMs that focus on students’ confidence judgements, and demonstrates how the OLM presents the information to align the computer or teacher’s judgement about the student’s answers with the students’ own judgement about the correctness of their answers. Further, it examines how the OLM is maintained when a misalignment is detected.

## **2.4 Aligning Multiple Belief Sources in Open Learner Models**

Representing student knowledge in the learner model can allow students to identify any strengths and weaknesses in their knowledge. The learner model uses evidence based on students’ answers to the system-automated assessment. However, students’ answers could reflect that the student either had mastered

the topic, or was simply guessing. Therefore, it is important that their self-perception is modelled separately. Providing students the opportunity to practise knowledge monitoring skills through visualising alignment may help them to identify their ability to assess their own knowledge.

Some OLMs promote learner reflection, by aligning confidence judgement with the teacher or computer evaluation. For example, in Reflect OLM (Kay, Li and Fekete, 2007), students' own judgement about the extent that they meet the teacher's criteria for answering problems are aligned with the teacher's criteria defined for each task. In Mr.Collins, CALMsystem, and ViSMod (Bull and Pain, 1995; Kerly, Hall and Bull, 2007; Zapata-Rivera and Greer, 2004), students' self-assessment of their confidence that the answers are correct is aligned with the computer evaluation of students' performance on the task.

In Reflect OLM (Kay, Li and Fekete, 2007), teachers create explicit criteria for each solution of each task to be consistent with the learning objectives of the task. Students are involved by making a judgement about how well they performed on a task with a set of marking criteria. The misalignment between the student and teacher's assessment represents how well the student understands the learning objectives. The learner model in Reflect OLM uses a text-based format to represent both student's beliefs of their judgement that they met the teacher's criteria for answering problems, and teacher's judgement, based on criteria defined for each task. Each judgement ranges along a four-point scale—poor, ok, good, and excellent (Kay, Li and Fekete, 2007). Allowing students to compare their own judgement to the teacher's judgement about their



performance can encourage them to think about their own self-rating of the assessment and promote reflection in action (Schon, 1983). An investigation by Kay, Li and Fekete, (2007) showed that although students learning was based on the teacher's worked examples, novice students were still over-rating their solution. The authors investigated the impact of the Reflect OLM on students' behaviour when using the system, conducting a qualitative analysis based on top and borderline students (Kay, Li and Fekete, 2007).

Where the Reflect OLM aligns both the extent that students met the teacher's criteria for answering problems and the teacher's criteria defined for each task, the following three systems (Mr.Collins, CALMsystem, and VisMod) provide examples that aligned the student's beliefs of the correctness of the answers given to that of the computer's evaluation about their performance. The learner model in Mr.Collins (Bull and Pain, 1995) represented the computer's evaluation of student performance and the student's confidence in their answers. The learner model in Mr.Collins used a text-based format to represent both the computer and the student's own beliefs. Each belief ranged along a four-point scale (very unsure, unsure, almost sure, and very sure). The student was then able to compare directly between the two beliefs and indicate if both beliefs are aligned or misaligned. In Mr.Collins, the learner model was maintained using a negotiation method, where both the computer and the student could initiate a discussion to resolve differences between the models. Such interactive maintenance could enable the system to develop a more accurate learner model.

Another example, the CALMsystem (Kerly, Ellis and Bull, 2008), was a negotiated learner model that used smiley faces to visualise the learner model that separately represented both the computer's evaluation and the student's confidence. Students could compare the two models and a negotiation could be initiated if one of the two parties (computer or student) disagreed with the content of the model. Both Mr.Collins (Bull and Pain, 1995) and CALMsystem (Kerly, Ellis and Bull, 2008) used two separated models to indicate whether the two models were aligned or not. Both systems studied the impact of the negotiation in order to have more accurate learner models and resolve any misalignment that may have arisen.

In ViSMod (Zapata-Rivera and Greer, 2004), the system employed conditional probabilities based on the strength of the relationships between nodes. Further, special aspects of learning (self-confidence, eagerness, assertiveness, and helpfulness) all influence the value and the colour of the node. The student could compare both nodes of the computer beliefs and their own beliefs, constructed on the basis of these learning aspects. Students could directly compare the two nodes to find misalignments based on the size and colour of the node. The system also combined both nodes to one combined node, also influenced by the aggregated value. The colour of the combined node comes from the node that most influenced its value. When the system detected a misalignment in the data sources for certain concepts, the system asked the learner to clarify these concepts to resolve the misalignment.

It is evident from these examples that both the computer and the student tried to reduce the misalignment between the computer's evaluation and student confidence using interactive maintenance mechanisms. The related studies did not investigate the impact of visualising the alignment using a controlled study, nor did the studies explore differences in how this information was represented within the OLM. The next section compares different visualisation within OLMs and students' perception about their preferred visualisation when more than one visualisation of the learner model is available.

## **2.5 Comparing Visualisations within OLMs**

Some OLMs show more than one representation of the learner model to students (Bull et al., 2016b; Conejo et al., 2014; Bull et al. 2014; Mabbott and Bull, 2004). These learner model visualisations can also vary how they present model information. For example, continuous or discrete scales can be used, with skill meters typically showing knowledge as a continuous variable through the degree of fill of the skill meter (Bull et al., 2016b).

Students tend to prefer bar graphs, skill meters, and tables to other visualisations such as smiley faces, word clouds, pictograms, grid plots, target plots, spider plots and tree maps (Law et al., 2015). Moreover, students often preferentially use the skill meter when other visualisation options are available (Bull et al., 2014; Duan, Mitrovic and Churcher, 2010). Keeping these preferences for chart type in mind, visualisations that include detailed information tend to be preferred to simpler visualisations. Thus, students prefer a skill meter that shows a fine-grained

model (Guerra et al., 2018) and prefer a detailed text-based representation to traditional bar graphs (Lazarinis and Retalis, 2007).

Preference for OLM representation is also influenced by the task being performed. Students prefer structured visualisations, such as graphs for pre-requisites and hierarchal trees, to support them when selecting their next task, but they prefer a skill meter visualisation for the purpose of comparison (Bull et al., 2016a).

Based on evidence in the literature about student's preference for the OLM representation, this thesis considers implementing the alignment between student's self-assessment of their confidence in the correctness of the answers given, and the system evaluation about their performance based on the representation of the skill meters that considered the most preferred and simple visualisation (Bull et al., 2014; Duan, Mitrovic and Churcher, 2010).

The next chapter discusses how visualising alignment is represented using two methods applied to the skill meter. The first method is the combined view, where data from two sources are combined using visual variables to indicate the alignment between students' correctness of the answer and students' confidence that the answers are correct. The combined view was implemented based on the proposed work of Demmans Epp and Bull (2015), where visual variables could indicate uncertainty information in OLMs, which could reduce the learner's cognitive load. The second method is the expanded view, where the two sources of data are represented separately using two skill meters. Although the combined view could allow students to have the overview picture without intensive attention, the

expanded view could let students spend more time aligning the two skill meters to indicate if a misalignment has occurred in their learner model. The expanded view has been used in prior OLMs, where students could compare their level of knowledge and their confidence judgement about their level of knowledge. For example, this occurs using text in Mr.Collins (Bull and Pain, 1995) or using smiley faces and text in CALMsystem (Kerly, Ellis and Bull, 2008). A pilot study was conducted to inform the design of the representation of the alignment information in OLMs, based on the two methods of the alignment applied to the skill meter, discussed in next chapter.

# **Chapter 3 : VISUALISING ALIGNMENT WITHIN OPEN LEARNER MODELS (OLMS)**

---

This chapter describes how information visualisation can play an important role in indicating alignment between two data sources within the underlying learner model in Open Learner Models (OLMs). In this thesis, alignment refers to correspondence between the students' own confidence in the correctness of the answers given and the computer's evaluation of the student performance.

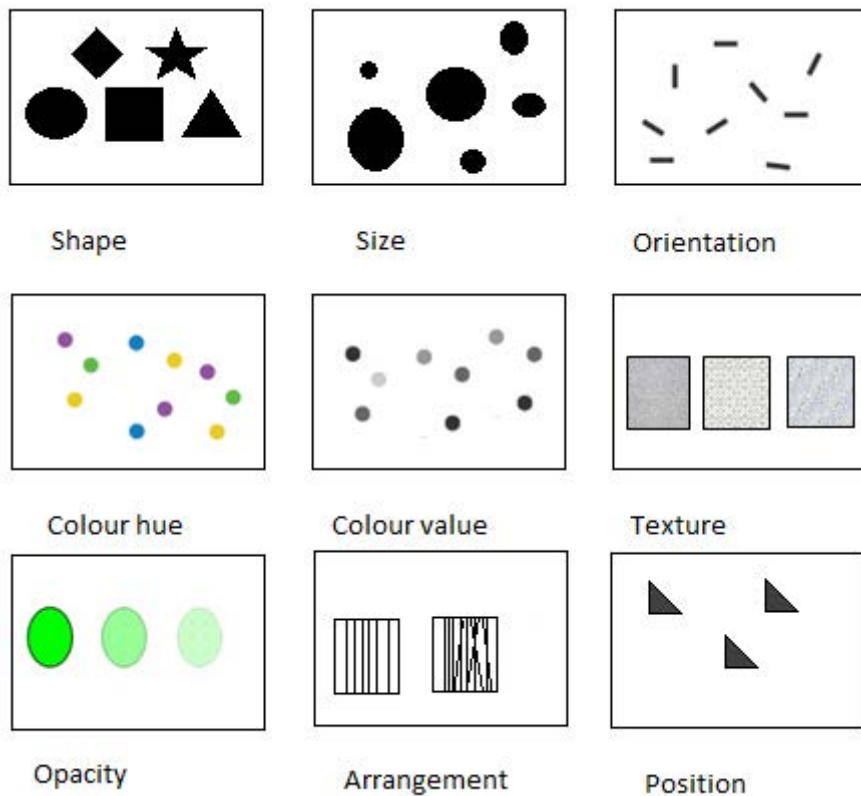
In this chapter, the issue of visualising alignment using the student's learner models is considered by providing different examples of how perceptual styles are applied in OLMs. It concludes with an overview of the system design and architecture of a currently existing system (OLMlets). The OLMlets system has been extended to add new features to support students' knowledge monitoring skills, by visualising alignment to fulfil the aim of the studies explored in this thesis.

## **3.1 Information Visualisation**

Information visualisation may include the quality of the data, such as the inconsistency, inaccuracy, incompleteness or errors associated with the data. All of these are considered types of uncertainties in the data (Correa, Chen and Ma, 2009). Specifically, this chapter considers how visual variables can be used to visualise inconsistencies (i.e. misalignment) that can be inferred from the

students' learner models (OLMs). Researchers argue that the use of well-selected variables to indicate inconsistency can aid users to understand the patterns represented by those variables without taking more time on the task (MacEachren, 1992).

Bertin (1967) identified seven visual variables: location, size, colour hue, colour value, grain, orientation, and shape. A further two visual variables have been suggested by Morison (1974), which are colour saturation and arrangement. MacEachren (1992) added another three visual variables: blur, opacity, and resolution. Visual variables can be used to maintain ease of comprehension for the data with the associated inconsistency (Zuk and Carpendale, 2006). Figure 3-1 presents some examples of visual variables.



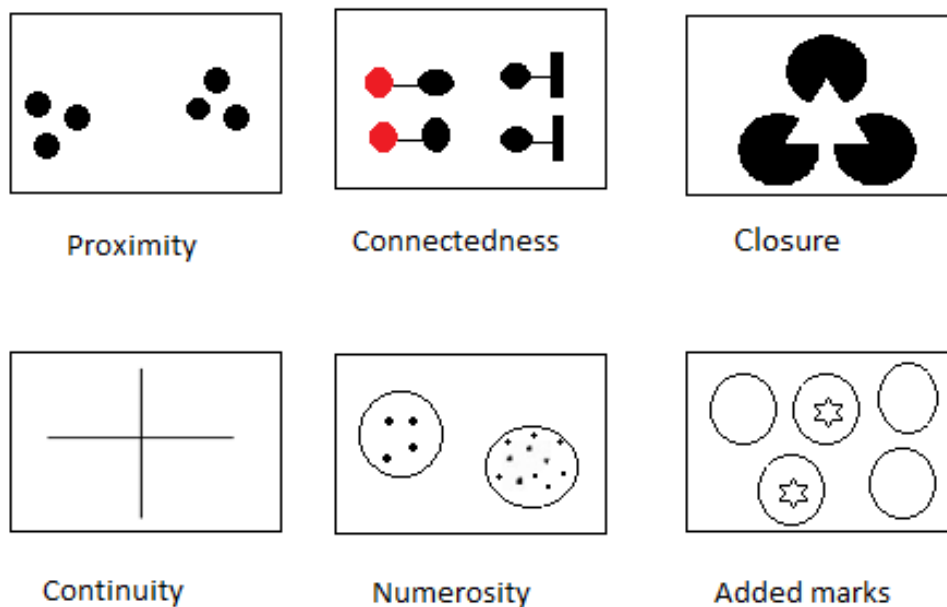
*Figure 3-1. Examples of visual variables.*

Researchers suggest that visual variables can be classified into two methods for indicating inconsistency or any other types of uncertainties in the data. The first method uses intrinsic techniques, such as the opacity of the colour, colour hue, colour value, and saturation. These intrinsic visual variables alter the existing display without adding objects. Conversely, extrinsic techniques add objects to the visual presentation to represent inconsistency, for example lines, glyphs, or dots (Kinkeldey, MacEachren and Schiewe, 2014).

Gestalt psychology introduced methods that aid people's ability to understand and interpret patterns depicted by visual variables (Ware, 2004). Some visual



variables have been shown to preattentively processed such as blur, added marks and numerosity. Gestalt psychology identifies people's responses to the different patterns of visual variables (Ware 2004), according to a set of basic principles, organised into five main categories such as Proximity, Similarity, Continuity, Closure and connectedness. These categories provide a framework by which it is possible to understand how people perceive the world, and how they respond to stimuli. This is applied specifically to the way that data is communicated, using visual patterns. Figure 3-2 presents examples of Gestalt principles.



*Figure 3-2. Examples of Gestalt principles.*

Proximity indicates that graphical items that are close together are perceived as grouped items, whereas connected items appear as a single item. The Gestalt principle of closure refers to an innate tendency to view incomplete items as

complete, closing any gaps between items to view asymmetric stimuli as symmetric. Lines or dots that continue in the same path tend to be viewed as a group. Numerosity refers to the perception of number within a grouping. For example, one single grouping of dots will appear greater than several smaller groupings with fewer dots each (where the same number of dots is used in both instances). Added Marks visual style allows people identify items or objects that contain the same mark to be classified as belonging to the same group (Ware 2004).

Bertin (1967) suggested that each of the visual variables has properties that may clarify certain information. Bertin classified the properties of the visual variables into four categories, as follows: orderable, associative, selective, and comparable (ibid.). Variables with properties that are orderable enable people to interpret the order of the items without being assigned an order. Associative variables represent the same variation as groups. Selective variables enable people to perceive visual style as an individual item. The final characteristic enables people to perceive the variable quantitatively, therefore being measurable in terms of being more or less, thus comparable. Table 3-1 presents an overview of the nine visual variable proposed by both Bertin (1967) and Gestalt (Ware 2004) discussed earlier, and their properties.

Table 3-1. Visual variables properties (Demmans Epp and Bull, 2015).

Properties	Visual Variables								
	Closure	Size	Orientation	Arrangement	Value	Opacity	Texture	Added Marks	Numerosity
Orderable	☒	☑	☒	☒	☑	☑	☒	☒	☑
Associative	⊙	☑	☑	☑	☑	☑	☑	☑	☑
Selective	☑	☑	☑	☑	☑	☑	☑	☑	☑
Comparable	☒	⊙	☒	☒	☒	☑	☒	☒	☑

☑ True, ☒ False, ⊙ Partially true

It is critical to understand how people can perceive and understand visual variables, when designing a system that uses visualisations, so that can people effectively interpret the data depicted by those visual variables. The next section describes how visual variables can be applied to different representations of OLMs to indicate misalignment or any other inconsistency inferred in the learner model.

### 3.2 Different Methods to Visualise Alignment in OLMs

Visualisation is an effective tool that could accurately communicate data. To fully understand the data, visualisation research has often considered including the error and uncertainty, which associated with the data (Bonneau et al., 2014). In OLMs, inconsistency could be considered a form of uncertainty, which could be

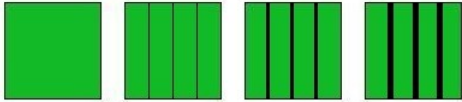
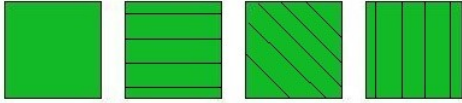
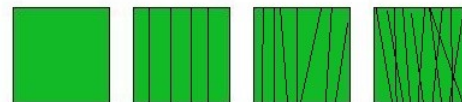
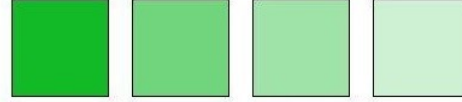
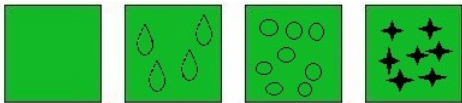
inferred when the learner model is constructed from multiple data sources, or when a conflict occurs between two data sources. For example, a student may perform well in one task and perform poorly in all other tasks. Further, there may be misalignment between students' confidence regarding the correctness of answers and students' performance on the actual answers.

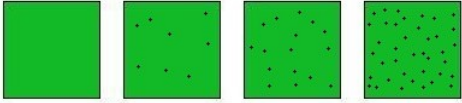
Demmans Epp and Bull (2015) proposed different methods to visualise inconsistent data in OLMs. Visualising the quality of these data could help students and teachers to understand the data better, and to make better decisions related to their learning. Visual variables can be manipulated to show different levels of variables, which can be used by system designers to show the level of the misalignment in the model. Some people might have difficulty in distinguishing between 10 levels of visual variables. Research shows that most people can effectively distinguish between four levels of each visual variable (aligned, low misalignment, moderate misalignment, and highly misalignment) (Boukhelifa et al., 2012). In Table 3-2, descriptions of the visual variables are presented, which represent the misalignment with the representation of the four levels applied in a skill meter visualisation in the exploratory study in Chapter 4. Skill meter was chosen because it was the most preferred visualisation when other visualisations were available (Bull et al., 2014; Duan, Mitrovic and Churcher, 2010).

In this section, examples of visual variables that are provided in Table 3-1 are applied to different OLM representations. These OLM representations provide examples to system designers when the domain includes levels of topics, and

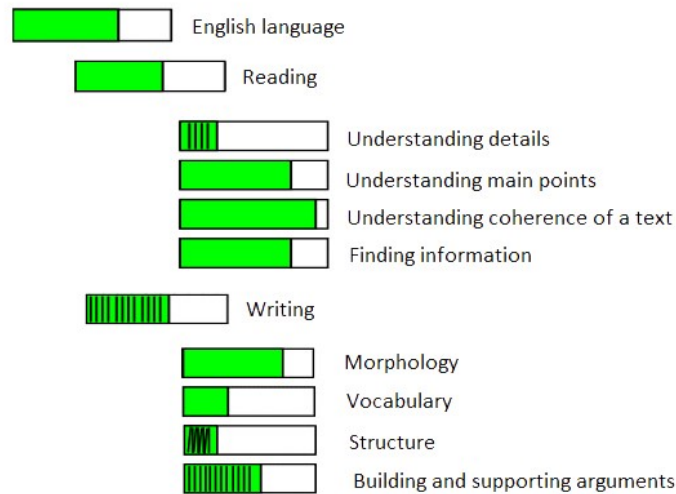
each topic has its own sub-topics. Topics within the learner model may be influenced by the level of the inconsistency of the sub-topics.

*Table 3-2. Descriptions of the visual variables with examples of levels of misalignment.*

Visual variable: Description	Levels of misalignment (from aligned to highly misalignment).
<p><i>Size</i>: the thickness of the line. The thicker the line the more misalignment.</p>	
<p><i>Orientation</i>: the rotation of the lines. The vertical the lines the more misalignment.</p>	
<p><i>Arrangement</i>: changes in the consistency of the lines. The messier the line the more misalignment.</p>	
<p><i>Opacity</i>: changes in the transparency of the colour. The more transparent the more misalignment.</p>	
<p><i>Added marks</i>: elements with the same mark belong to the same level of misalignment.</p>	

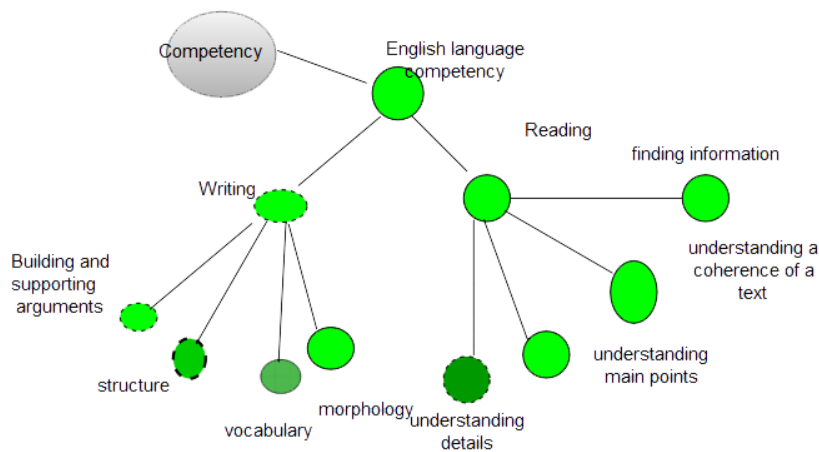
<p><i>Numerosity:</i> Change in the number of dots within an item tends to be belonged to the same group. The more dots the more misalignment.</p>	
--	--

The visualisations presented in this section are based on the OLM visualisation provided in Chapter 2 (Figure 2-1). The visualisations have been implemented in such a way that they all show the same content of the learner model. This enables the identification of how different visual variables can be used in the learner model using a variety of OLMs representations. Based on Gestalt principles (Ware 2004), people can perceive the same patterns as belonging to the same group, or the same level of misalignment when patterns are used to show the misalignment or inconsistency in the model. The levels of misalignment that are used in the following examples are based on the four levels presented in Table 3-2. For example, Figure 3-3 shows how arrangement could be applied in skill meters. The topic (structure) is represented with unorganised lines, which indicate high misalignment. In these examples, the level of misalignment in the sub-topics could inherit to the parent main topic, showing that the main parent topics include some misalignment in their sub-topics. This assumes that each sub-topic includes students' self-assessment that is compared with the computer evaluation of students' performance, and shows how both the computer evaluation and the students' confidence in the assessment are aligned.



*Figure 3-3. Arrangement is used in the skill meter to indicate topics that include inconsistent data (i.e. misalignment) in an OLM.*

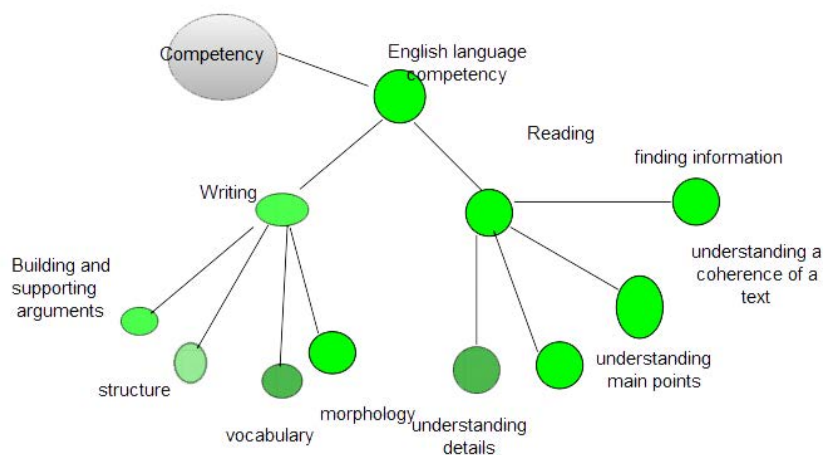
The level of knowledge in the network visualisation presented in Figure 3-4 is indicated by the size and colour. The larger and brighter the colour of the node, the higher the level of knowledge for the topic is achieved for the learner. The node could be indicated with a dashed line around the edge of the node to represent that there is a misalignment associated with that topic's assessment. The thicker the dashed line reflects the higher misalignment (Figure 3-4). Moreover, the misalignment could be inherited to the parent topic from its sub-topics (this can also be seen in the skill meter).



*Figure 3-4. The size of dashed line is used in a network to indicate topics that include inconsistent data (i.e. misalignment) in OLM.*

Figure 3-4 shows that the English language topic has two main sub-topics: reading and writing. In addition, the writing topic has several sub-topics, where a low level of misalignment is indicated with thin dashed line surrounding the node in the relevant sub-topic (building and supporting arguments). A high level of misalignment is represented in the other sub-topic (structure) indicated by the thicker dashed line around the node. The other two sub-topics show that the data evidence is aligned. This misalignment in the sub-topics influenced the parent topic (writing) that was visualised with a low level of misalignment. The level of misalignment of the parent topic is based on the average of misalignment levels of its sub-topics (one had low misalignment, one had moderate misalignment, and two topics were aligned). The opacity of the colour could be demonstrated to show the misalignment level in the node of the network visualisation (see Figure 3-5).





*Figure 3-5. Opacity is used in a network to indicate topics that include inconsistent data (i.e. misalignment) in an OLM.*

Some OLM representations may be more suitable than others based on the number of topics that the learner model has constructed. As has been discussed previously, a tree map could be useful when the domain includes a large number of topics (Brusilovsky, Hsiao and Folajimi, 2011; Bull et al., 2016a). In the field of simulation and visualisation, the brightness of the colour and line thickness has been used as a way to show data quality (Griethe and Schumann, 2006), such as misalignment in the context of this thesis. Following from these methods, misalignment could be indicated in the tree map using a dashed line around the topic border (Figure 3-6). The thickness of the borderline could reflect the level of the misalignment. This also could be designed using the opacity of the colour for each rectangle within the tree map, as shown in Figure 3-7.

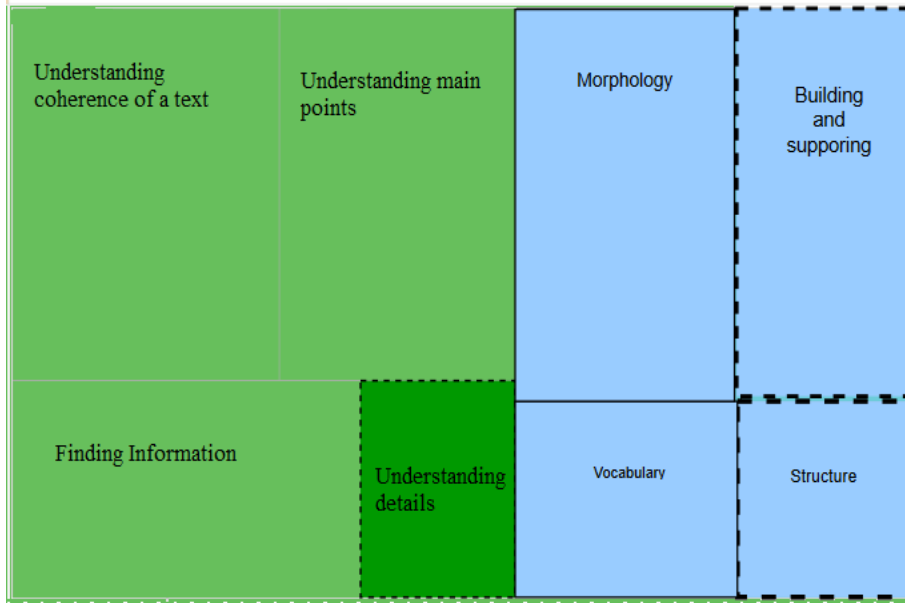


Figure 3-6. The size of dashed line is used in the tree map to indicate topics that include inconsistent data (i.e. misalignment) in an OLM.

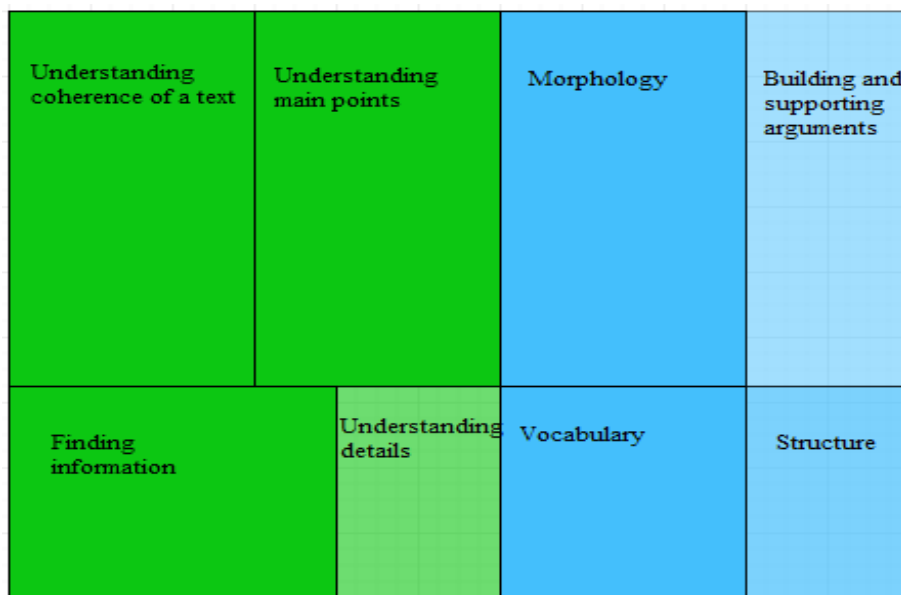


Figure 3-7. Opacity is used in the tree map to indicate topics that include inconsistent data (i.e. misalignment) in an OLM.

Another learner model visualisation is a word cloud, where the size of the text indicates the strength and weakness of the student's level of knowledge. Larger

text allows people to identify stronger topics quickly and smaller text shows that the student has some difficulty in the topic. Some OLMs separate stronger and weaker topics of the students' level of knowledge into two boxes, as shown in Chapter 2 in Figure 2-1 (Bull et al., 2016b). To display misalignment, text could be blurred incrementally. For example, the topic (structure) is fuzzier than the other topics in the word cloud, which represents a high level of misalignment within the topic. Grouping of sub-topics could be indicated with the use of colour to differentiate between each group. In Figure 3-8, two colours (orange and blue) are used to represent two groups. The use of colour indicates some structuring of the domain; otherwise, it is difficult to differentiate sub-topics that belong to the same parent topic with word clouds.



*Figure 3-8. Blur used in a word cloud to indicate topics that include inconsistent data (i.e. misalignment) in an OLM.*

As illustrated in Figure 3-9, radar plots can display the misalignment in the model by using a dashed line assigned to a topic that includes the associated misalignment data. This can be seen in above visualisation for the network and tree map representation.

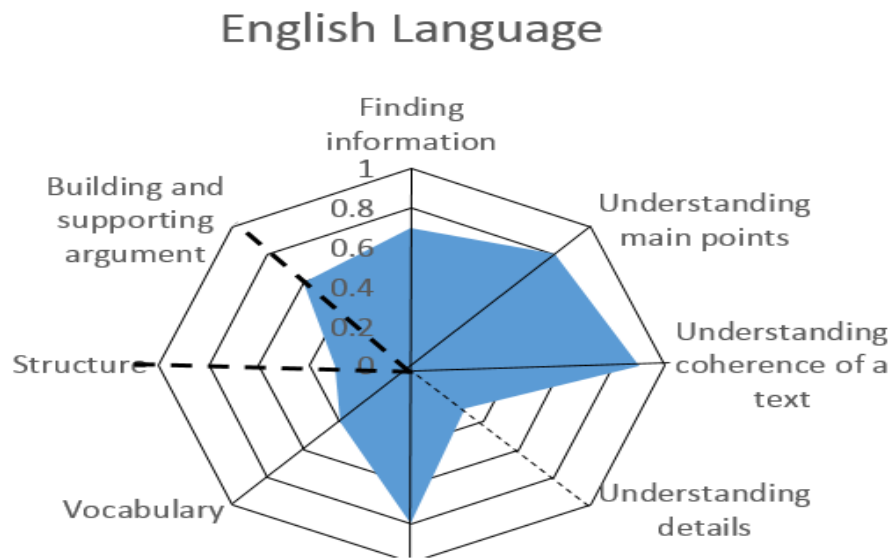


Figure 3-9. The size of the dashed lines is used in a radar plot to indicate misalignment in the OLM.

This section builds on the work of Demmans Epp and Bull (2015), which illustrated different methods that could be used to indicate inconsistency in the model within OLMs.

As indicated in Chapter 2 of this thesis, OLMs may promote students' metacognitive skills and can be used effectively to support students as independent life-long learners (Bull and Kay, 2013). Providing the learner with information about their confidence on the correctness of the answers and the computer evaluation on their performance through visualisation may promote students' motivation to

reveal the misalignment that occurred in topics of their learner model as a result. This enables students to understand their perception of both their confidence and their performance, which aids students to plan and select their learning strategy more effectively. In turn, this could promote their metacognitive skills (Bull and Kay, 2013).

The next section provides an overview of the system design and architecture of the existing system (i.e. OLMlets OLM), how the system was built to implement model alignment, and how it was represented in the learner model using graphical or text-based view.

### **3.3 System design and architecture (The OLMlets-OLM System)**

OLMlets is an educational technology that uses student answers (to pre-set questions) to provide adaptive feedback to students, based on the correctness of their answers. This correctness information is used to infer student knowledge, which then informs the representation shown to users. OLMlets does not adapt questions based on students' level of knowledge; rather it adapts the feedback they are given based on a learner model. OLMlets shows aspects of the underlying learner model to students so that they can monitor their knowledge (as measured by the correctness of the questions). This approach to adaptation is consistent with a recent review that argues presenting information about the correctness of student answers to questions is a core element of OLMs, which can affect student learning (Bodily et al., 2018).

OLMlets has been used extensively in real-class settings and is hosted by the School of Electronic, Electrical and Systems Engineering at the University of Birmingham. OLMlets was first used in 2006 (Bull and Mabbott, 2006) and has been used to support student learning across a variety of domains that include engineering (Bull and Gardner, 2010) and the mechanics of adaptive learning environments (Alotaibi and Bull, 2012). To fulfil the aims of this thesis, I have modified this established learner-modelling, educational system by extending its features to allow students to inspect the visualisation of the alignment between system assessments and student self-assessments.

### **3.3.1 The OLMlets interface**

OLMlets can be accessed through the URL <http://olmlets.bham.ac.uk>. After logging into the system, students can then select the course and topic that they want to study (Figure 3-10). Once students have selected a course, they can click on the Q button (located next to the tabs 'Revision1' and 'Revision2' in the lower window). This takes them to the systems' automated assessment of their knowledge for a specific topic. Students also can view the history of the last five questions and answers they provided by clicking the H button.

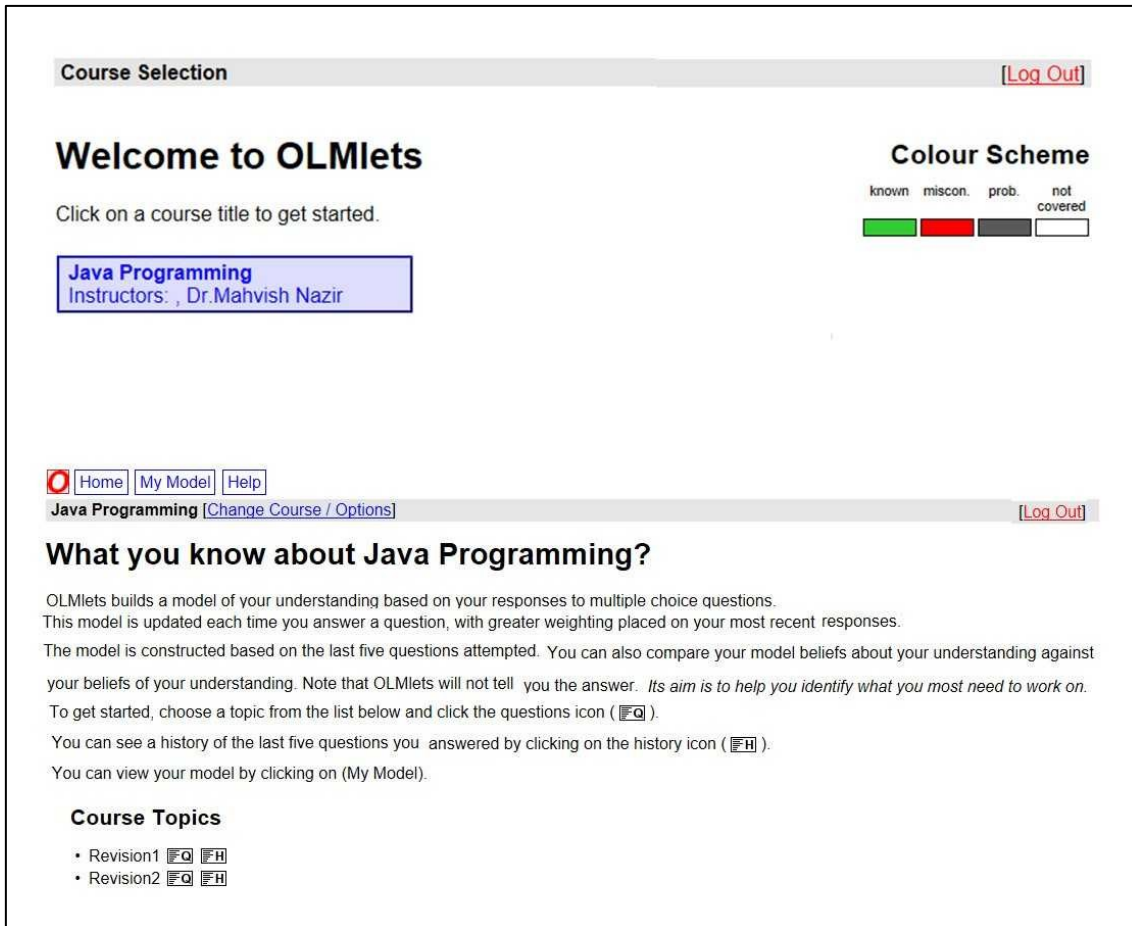


Figure 3-10. Screenshot of OLMlets interface: selecting the course (top), selecting the topic (bottom).

Each topic has a set of questions. Students are able to attempt questions more than once. However, every time the student completes the whole set of questions, the system provides a message to the student stating 'You have now answered all the questions on the topic. If you choose to proceed, you will be presented with questions you have seen before. If you do not wish to do this, choose another topic'.

### 3.3.2 OLMlets Base system

In a manner consistent with overlay models (Carr and Golstein, 1977), OLMlets uses answers to multiple-choice questions to create a model that represents students' weaknesses, strengths, or misconceptions (Bull, Jackson and Lancaster, 2010), as shown in Figure 3-11.

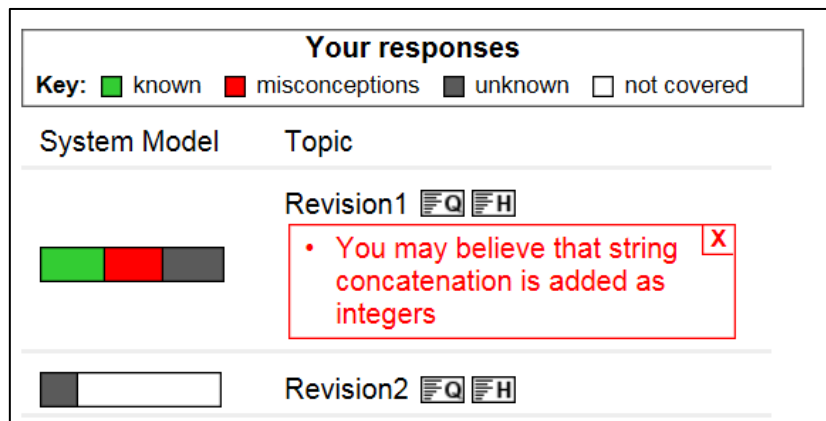


Figure 3-11. The standard skill meter within OLMlets showing details of a misconception in the course of Java Programming.

OLMlets has a misconception library that maps each individual question with the answer that provides evidence of a student misconception. These misconceptions are entered by instructors based on their experience and knowledge of the domain. Each question may or may not have a misconception associated with one of its answers.

OLMlets uses an algorithm that weights student responses based on when the question was answered; newer evidence is weighted more heavily and older evidence has less influence. OLMlets models students' level of knowledge based on the five most recent responses for each topic. This weighting mechanism keeps the learner model current, as long as the learner continues to use the



system actively. Other adaptive learning systems take into account additional parameters when weighting student performance, such as initial learning parameter, slip parameter, guess parameter, and learning parameter (Corbett and Anderson, 1995); however, they typically do not address issues of temporality in student modelling.

Questions in OLMlets are expected to have similar difficulty levels; thus, the weighting mechanism is not influenced by the difficulty of the question. However, the questions in OLMlets are not created following the more traditional ITS approach of deep domain modelling. In OLMlets, teachers are responsible for creating questions and associating them with knowledge components. This offers great flexibility, and makes porting the system from one educational domain to another relatively easy. In addition, it makes alignment of the system's questions with individual curricula relatively easy in comparison to many other ITSs. However, it can make modelling learner knowledge more difficult as it is unrealistic for instructors to identify question difficulty accurately, making the decision to exclude this as a modelling parameter reasonable, given the desire for a system that moves flexibly across educational domains.

Figure 3-12 presents a flow chart, showing how learner model values are calculated within the base system OLMlets (Bull and Mabbott, 2006). The flowchart shows the weighted calculation for a single question. Every time the student answers a question, the system retrieves the number of questions answered and divides them by 5 to calculate how many questions are remaining to fill the skill meter (number of questions = 5). It then updates the state of the learner model.

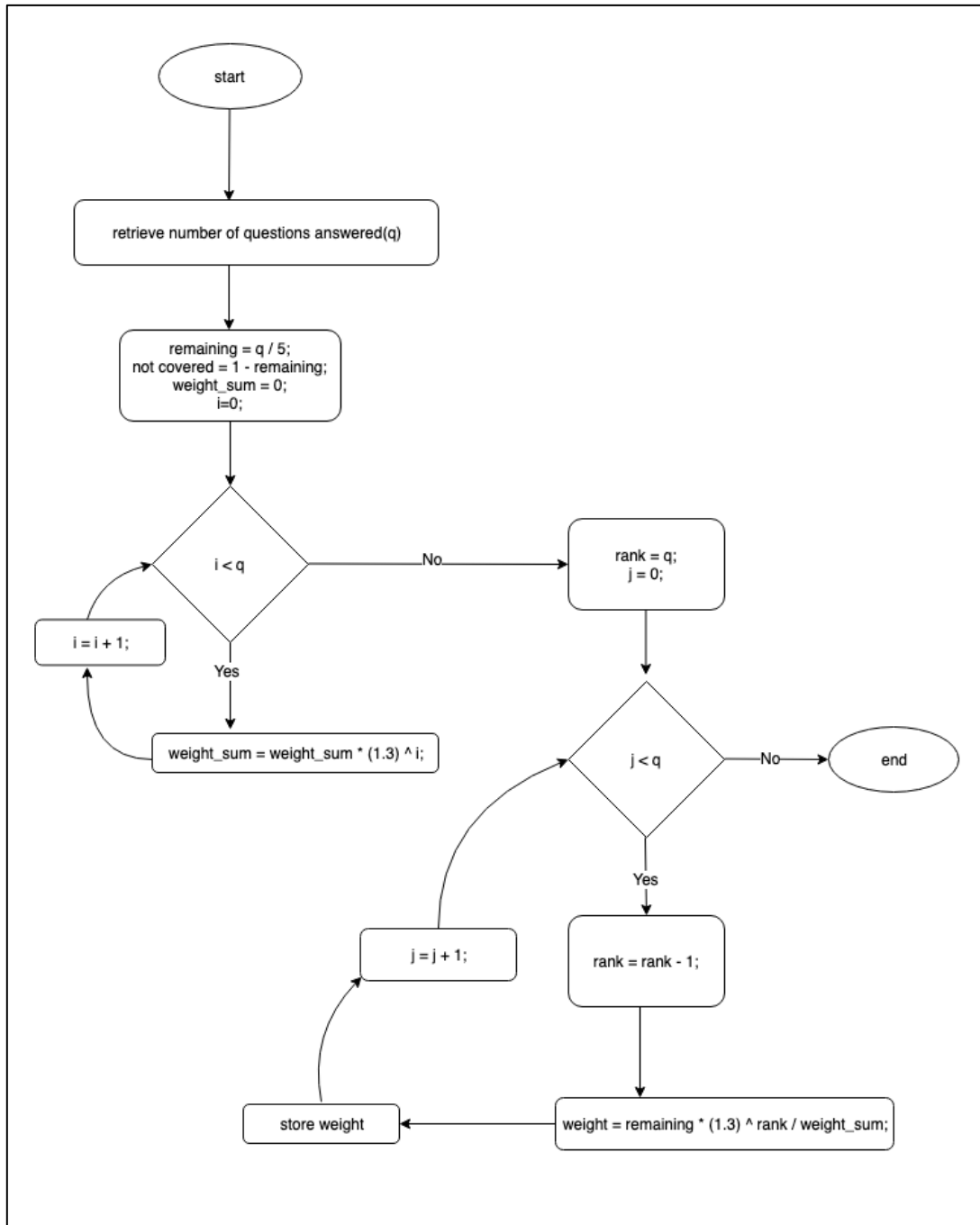


Figure 3-12. Flow chart of the calculation of the weighted mechanism for a single question.

OLMlets uses skill meters to communicate the underlying learner model to students, as shown in Figure 3-13. Initially, the skill meter is empty, showing that the

student has not yet answered the questions associated with that skill meter. When the student starts answering questions, the learner model represents the student level of knowledge based on the system's evaluation of their answers to questions. For example, in Figure 3-13, the student answered only one question in Revision 2, and performed incorrectly. Each segment of the skill meter is filled horizontally with a colour based on the accumulated weight for that segment while answering questions. The skill meter is filled completely when the sum of all three segments (known, misconceptions, and unknown) are equal to 1. The skill meter shows that a segment has been filled in grey based on how much weight the question holds, and keeps the remaining portion of the skill meter empty (not covered). A green segment within a skill meter indicates that the student answered the questions correctly. Students who answered the previous five questions correctly will have a skill meter that is completely filled in green, which is represented by the learner model as a score of 1.

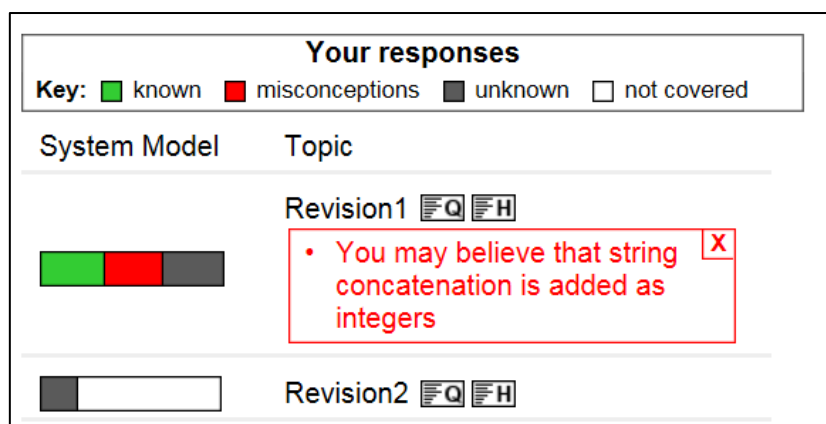


Figure 3-13. The standard skill meter within OLMlets showing misconception details in the Java Programming course.

If the student selects an answer that indicates a specific misconception from the misconception library, a red segment is filled in the skill meter and additional information is provided to students. This information is intended to explain the misconception held by the student. Misconceptions are incorrect answers that occur, due to a serious misunderstanding of a basic concept (Tasuoka, 1983). For example, Figure 3-13 shows the statement of the misconception in red 'You may believe that string concatenation is added as integers'. Seeing a specific misconception (for example, the concept of adding strings in a programming language in Figure 3-13), may encourage students to undertake more activities to understand the concept correctly. Here, misconceptions are not a focus in this thesis, as all studies were conducted in real class settings, and instructors were not concerned about misconceptions. Grey segments show that the student has provided incorrect answers (unknown in Figure 3-13).

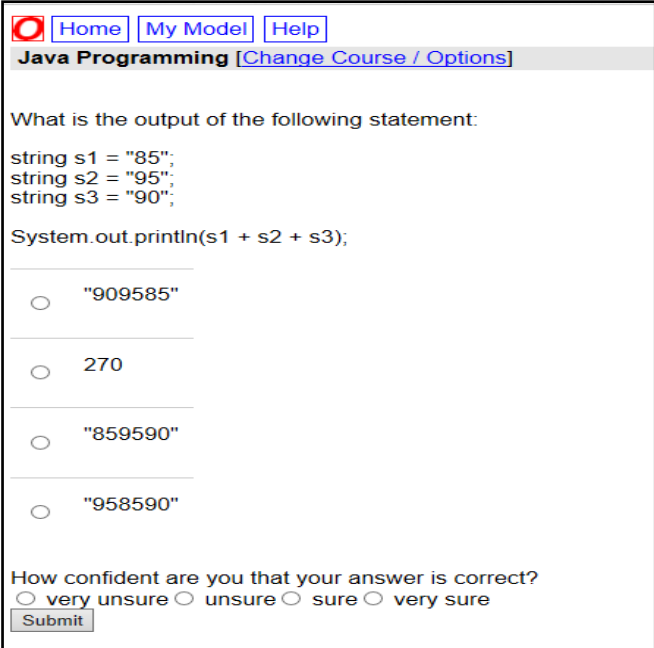
OLMlets has additional features (not used in this thesis), such as giving students the ability to release their own learner model to their peers, and enabling them to compare their level of knowledge to the instructors expected knowledge to that of the student. Students can also compare their level of knowledge to that of the whole class who are registered in that particular module.

For this thesis, OLMlets was used as individual OLM model to ensure the studies were conducted in controlled conditions. It was important that there were no influences from other metacognitive behaviour, such as collaboration with peers and aligning their performance with the instructor's expected beliefs of the student knowledge, that might impact on student confidence judgement in the

correctness of the answers. To investigate the impact of the alignment between the students' performance based on the computer evaluation and the students' own confidence in their answers (which would be given without external influences), other features were extended in OLMlets. The next section details the extended feature of model alignment in OLMlets.

### 3.3.3 Modelling alignment

OLMlets has been modified to allow students to assess their confidence in the correctness of their answer to each question, using a 4-point Likert scale: very sure, sure, unsure, and very unsure. Students can inspect their learner model by clicking 'My Model' at any time while answering the questions, as shown in Figure 3-14.



The screenshot shows a web interface for a Java programming question. At the top, there are navigation buttons: a red circle icon, 'Home', 'My Model', and 'Help'. Below this is a header bar with 'Java Programming' and a link '[Change Course / Options]'. The main content area contains the question: 'What is the output of the following statement:'. The code snippet is: 

```
string s1 = "85";  
string s2 = "95";  
string s3 = "90";  
  
System.out.println(s1 + s2 + s3);
```

 Below the code are four radio button options:   
 "909585"  
 270  
 "859590"  
 "958590"  
At the bottom, there is a self-assessment question: 'How confident are you that your answer is correct?' with four radio button options:   
 very unsure  unsure  sure  very sure  
A 'Submit' button is located at the bottom left of the form.

Figure 3-14. An example of a question and a self-assessment in the extended OLMlets

The system evaluates students' answers based on the correctness of the answers, and students assess their confidence of having entered a correct answer for that question. To construct the learner model, the system models the student level of knowledge as a continuous value between 0 and 1. A value of 0 indicates that the student currently has no knowledge, and a value of 1 represents full knowledge, or mastery. Students' assessment of their own confidence is also represented as a continuous value between 0 and 1. The confidence model uses the same weighting algorithm as the computer evaluation.

The self-confidence options 'very unsure' and 'unsure', both represent that the student is unconfident that the answer is correct. When the student selects the confidence option 'very unsure', this is weighted more heavily; it accounts for the full weight of the question (100%). If 'unsure' is selected, the confidence level will be estimated using 75% of the weighting as unconfident and 25% of the weighting as confident. Similarly, the self-confidence option 'very sure' represents that the student is 100% confident. This option is therefore weighted as such. The self-confidence option 'sure' is represented as 75% confident and 25% unconfident.

The alignment between the two models is represented as a level where alignment values range between 0 and 1. Model alignment has four levels that show the degree of alignment: low (when the value of model alignment lies between 0.0 and 0.3), medium (when the value lies between 0.3 and 0.7), high (when the value is between 0.7 and 1.0), and aligned (when the value equals 1.0).

Misalignment arises in the learner model when system evaluation and student confidence conflict. For example, a student can select 'very unsure' as a

confidence level when they has answered correctly, which means that the system believes the student knows the answer. This indicates that the system's evaluation of student performance on questions answered is not aligned with the student's confidence that the answers are correct. Furthermore, if students answered the question incorrectly, and selected 'very sure', this results in a misalignment that is weighted more heavily.

### **3.3.4. Visualising alignment**

The alignment enables students to understand both their self-perceptions and their performance. While it is possible that a lack of confidence indicates that a student guessed their answer, it may simply indicate that a student does not believe in their abilities. Therefore, in the absence of additional information, it would be difficult to integrate student self-perceptions reliably into the assessment of their knowledge. Two methods were implemented in OLMlets to visualise alignment, based on the skill meter visualisation. The first method was the combined view, where the opacity of the colour is used to indicate alignment. The second method was the separated method where two skill meters were used to compare between the system's evaluation of student's performance and the student's confidence about the correctness of the answers.

#### ***3.3.4.1 Representation of alignment using the combined view***

The combined view (Figure 3-15) uses the opacity of the colour to indicate alignment, whereby the higher the transparency of the colour, the more misalignment

occurred in the learner model. The opaque colour represents that the system's evaluation of the answers being correct is aligned with the student's confidence about the correctness of the answers.

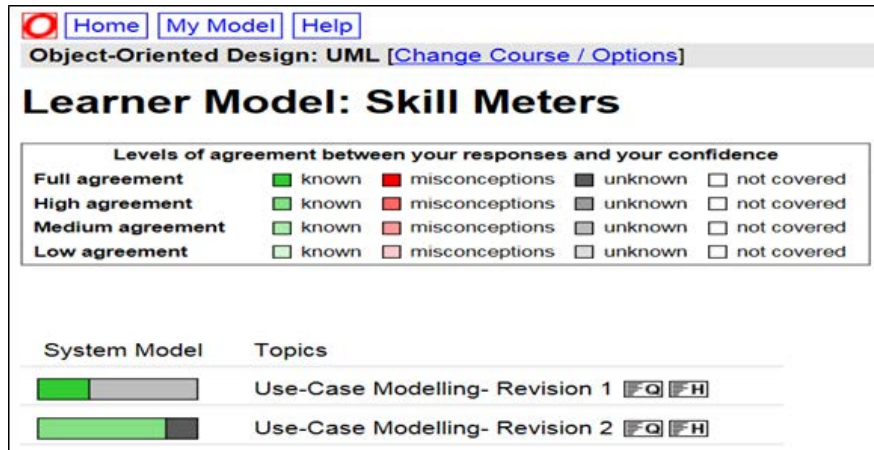


Figure 3-15. Combined view.

Each level of model alignment is represented by the opacity of the colour, with each level having a certain degree of colour transparency that indicates the level of alignment, as follows: low (alpha = 0.2), medium (alpha = 0.4), high (alpha = 0.6) and aligned (alpha = 1, fully opaque). The system represents the alignment (i.e., agreement) in the learner model using four levels, because according to Boukhelifa et al., (2012), most people can effectively distinguish between four levels of opacity, as shown in the skill meter legend (Figure 3-15).

In Figure 3-15, the first skill meter of the topic 'Use-Case Modelling- Revision 1' shows the transparency of the grey colour, indicating that the student answered the questions incorrectly while they were confident that the answer was correct; this indicates that the student was over-confident. In the second topic 'Use-Case



Modelling- Revision 2', the transparency of the green colour shows that the student was under-confident where the answers were actually correct.

### 3.3.4.2 Representation of alignment using separated models

The second method is the use of two skill meters presented separately. The first skill meter 'Your confidence' is represented by confident (green) and un-confident (grey). The second skill meter is 'your knowledge' shows the system evaluation of the correctness of the answers based on the colours of the original skill meter. Students can compare directly between the correct answer, represented by the green colour in the system's model 'your knowledge', with the green colour represented by confidence from the second skill meter 'your confidence' (Figure 3-16). Similarly, students can compare their lack of knowledge represented by the system using the colour grey in 'your knowledge' with the un-confident represented by the grey colour in 'your confidence'. Thus, students can compare directly between the two colours from the two skill meters to explore the alignment.

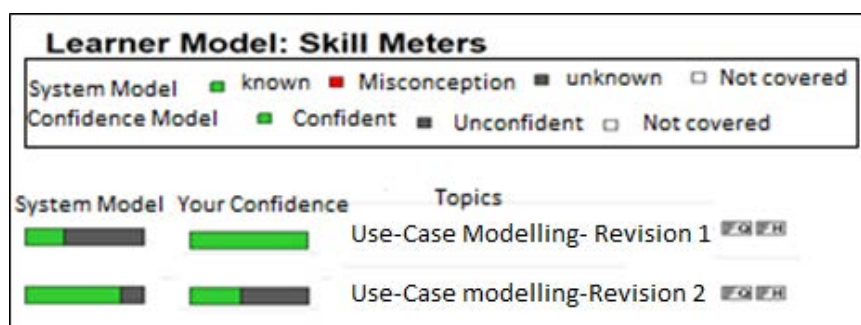


Figure 3-16. visualising alignment using separated models.

### **3.3.5 A pilot study for an alignment representation decision**

A pilot study was conducted to help to inform the design of the alignment information applied to the skill meter for students using two methods: combined or separated skill meters. The participants were 20 undergraduate students from the College of Applied Sciences in Oman. The students, all of whom volunteered to participate in the study, used the OLMlets system for one week. Students were randomly assigned to three groups: one control condition and two treatment condition groups. The control condition group used the original skill meter of the OLMlets base system (see Figure 3-13). The first treatment condition group used the combined view (see Figure 3-15) and the second treatment condition group used the separated model (see Figure 3-16).

From the 20 participants, only 10 students used the system and had valid logs (view the model >2 and answered question > 10) (original skill meter n = 5, combined view n = 3, expandable model n = 2).

The interaction between the participants and the system has been analysed to determine the pattern of users who were willing to look at their model, depending on their group. Two viewing behaviours of students were observed, whereby some participants looked at their model after almost every question answered, whereas others looked at their models at regular time intervals. For instance, two users viewed the model at two intervals: at 50% and 100% of the time taken to answer the questions. Another user viewed the model seven times, with a difference of around 10% between each view and with a total number of 57 questions

answered. Figure 3-17 shows the number of users in each group with the different patterns of viewing the model.

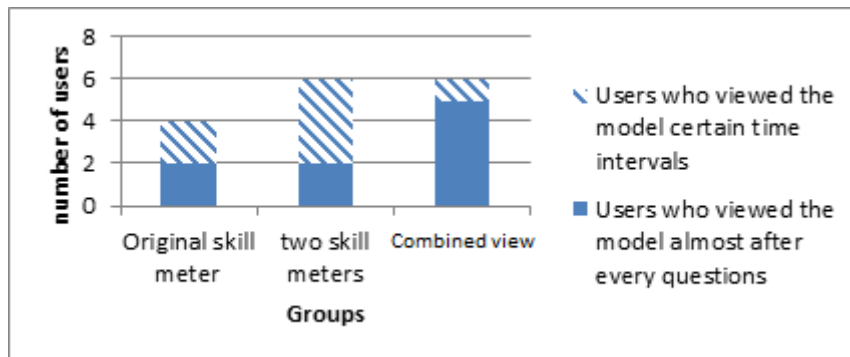


Figure 3-17. Patterns of viewing the model based on the number of users in each group.

Results showed that more students from the combined view were encouraged to view their learner model, almost after each question answered, whereas students from the separated models (Figure 3-15) had the higher number of questions answered (average = 116 questions). Students from the combined view had an average of 86 questions answered, whereas students from the original skill meter view had the least number of questions answered (average = 57). The combined model may have motivated students to align their confidence to their performance to reach the opaque colour, which might indicate an influence on students' behaviour to view the model.

The finding shows that both treatment conditions may have an influence on students' behaviour in using the system. This finding had an impact on my decision regarding how alignment is represented by building a representation that includes both views (combined and separated models) as one form of visualisation, called

the expandable model, as presented in Figure 3-18. Students in the expandable model can view the model using the combined version, or click to expand the models to two skill meters by clicking 'show models' (see Figure 3-18). The separated skill meters in the expandable model view were modified to be vertically aligned for easy comparison between the two skill meters.

Allowing students to have an option to view the separated models was based on the finding of the pilot study, where students who were in the combined model were more motivated to view the model almost after every question. The separated model could help students with difficulty in interpreting visual variables to have clear comparison between students' confidence in the correctness of the answers and the computer evaluation of students' performance on answers given. However, this design decision may hide some information that might be of interest to students if they were not aware of an option to expand the model, or they did not want to give additional effort by clicking to expand the model, which may influence the finding. However, before each study, I gave students a 15 minutes presentation to explain the system design and its features so students could benefit from the features included in the system. Further, having the option to expand the model would allow me to explore whether students were motivated to engage greater effort to get additional information for future design decisions. This aspect is explored in the study presented in Chapter 5.

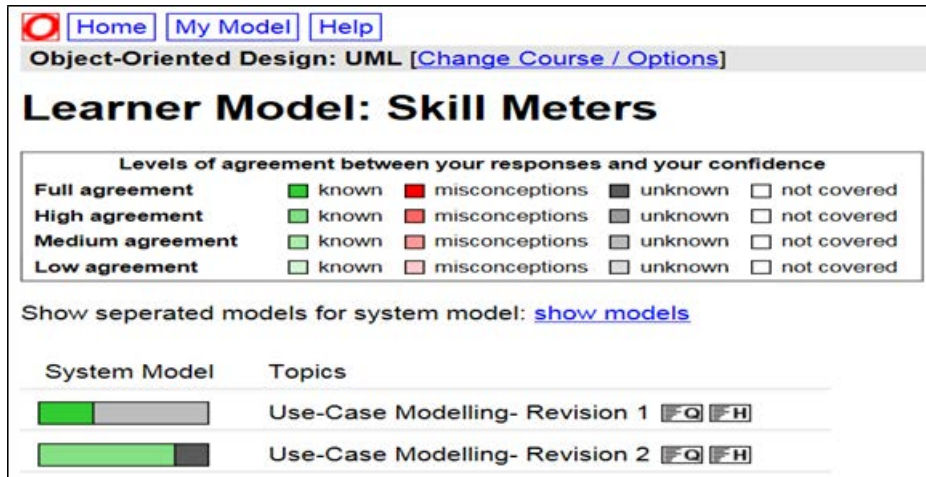


Figure 3-18. Expandable model, where the models are combined.

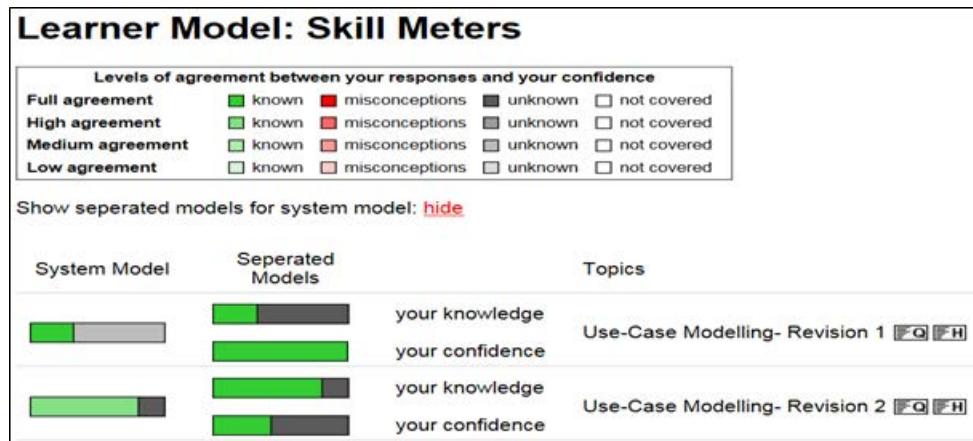


Figure 3-19. Expandable model where models are expanded.

The next section describes how the learner model is affected while the student engages in answering the questions.

### 3.3.6 The impact of students' answers on the learner model.

Table 3-3 presents the calculation after a student answered seven questions. For each answered question, the weight for each segment of the learner model (known, unknown, or misconception) is filled according to the calculated weight. The alignment value is initially assigned to 1 and decreases every time a

misalignment is detected by the weight of the question. For example, when the student answers the first question, the weight calculated (based on the flowchart in Figure 3-12) is 0.2. The skill meter is filled completely when the total sum for each segment (known, misconception, and unknown) is 1.

*Table 3-3. The system calculation based on the weighted questions of seven questions answered.*

Questions attempted	System evaluation			Student confidence		Alignment value			Student confidence option	Correctness of answers
	Known	Unknown	Misconception	Confidence	Un confidence	Known	Unknown	misconception		
<b>1</b>	.2	-	-	.05	.15	.85	1	1	“unsure”	correct
<b>2</b>	.17	.23	-	.21	.19	.93	.87	1	“sure”	incorrect
<b>3</b>	.40	.20	-	.24	.35	.75	.95	1	“unsure”	correct
<b>4</b>	.63	.17	-	.21	.59	<b>.50</b>	.96	1	“very unsure”	correct
<b>5</b>	.86	.14	-	.26	.74	<b>.33</b>	1	1	“unsure”	correct
<b>6</b>	.89	.11	-	.42	.58	<b>.55</b>	1	1	“sure”	correct
<b>7</b>	1	-	-	0.57	0.43	<b>.63</b>	1	1	“very sure”	correct

Figure 3-20 shows that the skill meter is affected as the student answers questions, based on student’s assessment of their confidence that the answer is correct and the system evaluation of student’s performance of the question, shown in Table 3-3. Figure 3-20(a) is filled with a portion of 0.2 in the colour green, showing that the student answered the first question correctly. In this example, the student selected ‘unsure’ as the self-assessment option, which is calculated as 75% of the weighted question as un-confident and 25% of the weighted question as confident. In Figure 3-20(a), the combined view shows that the skill meter

is filled with green (0.2 known, see Table 3-3). In the expanded view, the system evaluation (your knowledge) is filled by 0.2 with green and the student's own confidence is filled with 0.15 un-confident (grey -  $0.2 * 0.75$ ) and .05 confident (green -  $0.2 * 0.25$ ). In this example, there was a misalignment between the system evaluation about the correctness of the answer and the student's own confidence that the answer is correct. Thus, the alignment value was reduced for the known concept by the weight of the student's un-confidence (.15), which resulted in an alignment value of (.85). The alignment value affects the transparency of the colour in the combined view. The lower the alignment value, the higher the transparency of the colour and the higher the misalignment.

When the student answered the second question, the newer question had more weight: the weight of the second question was calculated as 0.23 and the first question's weight was reduced from 0.2 to 0.17. Figure 3-20(b) represents the student model when two questions were answered. The second question was answered incorrectly, while the student assessed their confidence as 'sure'. The student answered the next five questions correctly; however, student confidence that the answers were correct ranged between very unsure to very sure (Figure 3-20(c) – Figure 3-20(g)). Figure 3-20(g) shows the last five questions answered where the first two answers are now not included in the learner model.

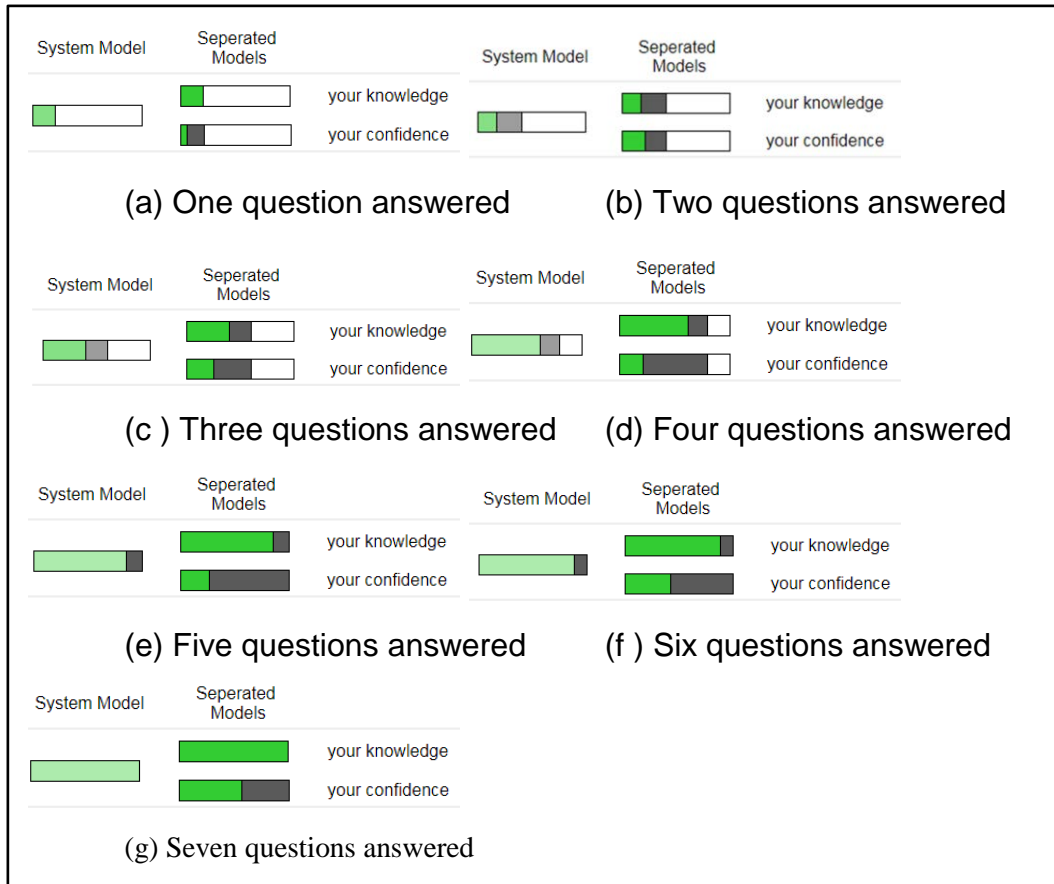


Figure 3-20. Shows how the skill meter (expandable model) is affected while the student answered seven questions based on Table 3-3.

To answer the thesis research question (How is the information's presentation format important to the alignment information in OLM?). OLMlets was extended to include two forms of alignment representation: the first representation is the graphical view discussed above and the second representation is the text-based representation, which will be discussed in the next sub-section.

### 3.3.7 Representation of alignment using text-based

OLMlets was extended to provide a text-based learner model that represents how the computer's evaluation of student answers is aligned with student beliefs about



confidence in the answers given (Figure 3-21). The modelling process used in the expandable model and the text-based learner model was the same. However, the text-based learner model converts the level of knowledge into four categories: low, moderate, high, and very high. The usage of the levels in textual description has been utilised in some prior OLMs (Bull, Pain and Brna, 1995; Kay, Li and Fekete, 2007).

The screenshot shows a web interface for 'Object-Oriented Design: UML'. It features a navigation bar with 'Home', 'My Model', and 'Help' buttons. Below the navigation bar, the title 'Learner Model: Text' is displayed. A legend defines the levels of system and user beliefs: 'Levels of system beliefs about my knowledge' and 'Levels of my beliefs about my knowledge', both with categories: low, moderate, high, and very high. The main content is a table with columns for 'Topics', 'Topic Covered', 'System's Beliefs' (known concept, unknown Concept), and 'My Beliefs' (known concept, unknown concept). Two rows of data are shown for 'Use-Case Modelling- Revision 1' and 'Use-Case Modelling- Revision 2'.

Topics	Topic Covered	System's Beliefs		My Beliefs	
		known concept	unknown Concept	known concept	unknown concept
Use-Case Modelling- Revision 1	100%	moderate	moderate	very high	-
Use-Case Modelling- Revision 2	100%	high	low	moderate	moderate

Figure 3-21. Text-based learner model within the OLMlets shows both (Computer's Beliefs, My Beliefs) specified in Study (Chapter 7).

This was conducted because providing many levels to represent the level of knowledge may cause confusion to the student and discourage them from using the OLM. As shown in the expandable model (Figure 3-18), the skill meter is completely filled when the student has answered at least five questions, which is also indicated by achieving 100% 'Topic covered' in the text-based model. Figure 3-18 and Figure 3-21 show two different representations of the learner model that include the same content of information.

The system shows a level for each 'known concept' or 'unknown concept'. The learner model uses a continuous score from 0 to 1. The system shows the "-" for

known or unknown concepts to indicate that no knowledge has yet been detected, or the value of the model is 0 for that particular concept. The system measures a low level of knowledge when the value ranges from 0.0 to 0.3; a moderate level of knowledge when the score is between 0.3 and 0.7; and a high level of knowledge is represented when the score is between 0.7 and 1 exclusively. The system shows a very high level of knowledge when the model score reaches 1. A very high level assigned for the unknown concept means that the students has very high difficulty level on the task. Students' confidence is represented as 'My beliefs' and the level of confidence is represented similarly as the level of knowledge. Each time the student answers a question, the system updates the model and specifies which level the student has reached. It is possible that the student might stay in the same level if the model score is within the range of that level. Students can compare between both beliefs to detect any misalignment between them. Figure 3-21 shows the text-based version for one of the student's learner models from a study from Chapter 7 of this thesis. The first topic in the learner model shows that the system assigns 'moderate' as the student's known concept, where the student feels over-confidence. This shows that both beliefs are not aligned. The second topic the student has a high level of knowledge based on the computer evaluation; however, the student's confidence shows as moderate level indicating that the student was under confident. This can reflect student awareness and encourage students to align both beliefs and reduce any misalignment detected in the learner model by answering more questions with more accurate confidence judgements.

While there are several forms of graphical design to display alignment, and while this thesis relies on two approaches, it would be beneficial to ensure that these are the most appropriate. The next chapter considers alternative visualisations to select the most appropriate, based on students' preferred method.

# Chapter 4 : PREFERRED METHOD OF VISUALISING ALIGNMENT IN OLM (EXPLORATORY STUDY)

---

## 4.1 Overview

The study described in this chapter explores students' preference for visualisation of the alignment between students' self-confidence assessment in the correctness of their answers and the computer's evaluation about the student's performance on answers given. The study answers the following research question:

1. What is the preferred method to visualise alignment for OLM users?

Based on the theory discussed in the introduction to this thesis, students are motivated to resolve misalignment when they are provided with the alignment information between their performance and their confidence in the correctness of the answers given (Kulhavy and Stock, 1989). An exploratory study showed that students were more active when viewing the learner model for their preferred representation when multiple views are provided (Bull and Mabbott, 2006). Some researchers suggest that participants prefer a method for indicating inconsistency, such as a dashed line, because of its noticeability among blur, grey scale, and sketchiness (Boukhelifa et al., 2012). However, prior research showed that participants correctly inferred information related to data quality (such as inconsistency) using their preferred method (Gerharz and Pebesma 2009). Research

has also indicated that lack of knowledge monitoring skills is attributed to motivation (Zimmerman, 1990). Therefore, it is important to explore the preferred method (visual variable) to visualise alignment that can encourage students to use the OLM. Thus, they will be able to interpret the information depicted by their preferred visual variable accurately, promoting their metacognitive skills.

Exploring students' preferred method to indicate inconsistent data (misalignment), has been explored outside the educational domain. In this study, a survey-based analysis to explore students preferred technique of visualising alignment is explored. This study is a novel study as, to our knowledge, no study explored students preferred method of visual variable that could be used to indicate any form of variability in the model in the field of OLMs and ITSs.

Six visual variables were chosen to be applied to a learner model visualisation (skill meter). Paired comparison questionnaires (including 15 pairwise comparison items) were printed in colour and distributed among students. Then, students judged the preferred visual variable in each pair. The collected data from this exploration study were analysed using two methods. The first method used the paired comparison method (Thurstone, 1927) which was the first scientific approach to measure pairwise comparison. For a stronger argument, the statistical method (widely used in research) was used. Exploring student's preferred method to visualise alignment in OLMs can provides hints about which approaches students may use, so further approaches can be considered.

## **4.2 Visual variables applied in OLM.**

Assessing the use of visual variables has been performed in other disciplines, to indicate inconsistency and other forms of uncertainty (Slocum et al., 2003; Thomson et al., 2005); however, little educational research has explored how to indicate and communicate inconsistency in the model, or other forms of variability in the model. Some systems have used different techniques of visual variables to communicate inconsistency to the user easily. For example, in OLMlets OLM (Bull, Jackson and Lancaster, 2010), the system uses the colour value to show the student's weaknesses, misconceptions, and strength in their level of knowledge, which is then compared to the instructor's expected knowledge of the student. Students can easily compare the fill colour of both skill meters for an overview of the inconsistency between their level of knowledge and what is expected of them. The visual variables of size and colour have been used in VisMod (Zapata-Rivera and Greer, 2004), where the system indicates the level of knowledge by using the size and the colour of the node.

Uther (2001) implemented a web-based visualisation of large user models (VIUM), whereby users were enabled to identify the outliers or interesting parts of the model quickly, using the position in the model. In addition, Uther (2001) suggested that opacity could be helpful for indicating components that reach or approach the value of interest to the user; components with values further away would fade to be more transparent or more opaque. This can help the user to identify the topics which they have high level of knowledge when the learner model contains a larger number of topics.

### **4.3 Visual variables selected for this study**

Demmans Epp and Bull (2015) proposed different techniques using visual variables to indicate any form of variability in the learner model that can be used in educational reporting. 50 visual variables were analysed, based on communication means, to depict uncertainty information (Demmans Epp and Bull, 2015). In this thesis, six visual variables were selected from those 50 variables, to explore students preferred technique to visualise alignment in OLMs (Demmans Epp and Bull, 2015). These six visual variables were then applied to skill meter visualisation used in the OLMlets system (discussed in Chapter 3). The skill meter was used because it was the most preferred visualisation for students when other visualisations were available (Bull et al., 2014; Duan, Mitrovic and Churcher, 2010). Table 4-1 shows the visual variables selected in this study, the associated percentages used for the visual variable to communicate inconsistent information, and the criteria used to select the visual variable. From the table, it can be seen that two variables were selected (based on 100% communicating uncertainty). The other two visual variables (numerosity and opacity) were based on the criteria of (> 80% and <100%) to communicate uncertainty. The last two variables were selected based on the criteria (>50% and < 80%).

Based on the types of visual variables discussed in Chapter 3, the visual variables listed in Table 4-1 can add objects such as lines, dots, or glyphs (i.e. extrinsic techniques); except for the visual variable (opacity), which is an intrinsic technique that can alter the existing display. Although saturation and colour hue are intrinsic techniques of visual variables and were allocated in the range above

50% and less than 80%, these variables were difficult to implement in the skill meter, because of the different colours used to indicate the students' level of knowledge. The visual variables selected should fit the design of the skill meter visualisation, which is the base visualisation used to implement the following visual variables to visualise misalignment.

*Table 4-1. Percentage of the use of different methods to represent only uncertainty according to the summary of 50 visualisations from different fields (Demmans Epp and Bull, 2015).*

<b>Visual variable</b>	<b>Uncertainty Only %</b>	<b>Selected criteria</b>
Added marks	100	100%
Arrangement	100	100%
Numerosity	88.9	80 - <100%
Opacity	83.3	80 - <100%
Orientation	72.7	50 - <80%
Size	52.0	50 - <80%

All six visual variables presented in Table 4-1 have selective and associative properties, based on Bertin's classification of visual variables reported by the analysis of Demmans Epp and Bull (2015). The associative property of the visual variable enables students to identify that each pattern of the visual variable belongs to the same group or level of alignment, which can easily provide the overview picture of the inconsistent data in the model. It also has the feature of selective variables, which allows students to classify each level of the visual variable as a single item.



#### **4.4 Designing the visual variables into OLM visualisation (skill meter)**

Based on the literature review of information visualisation discussed in Chapter 3, people can effectively interpret the level depicted with the visual variable for up to four levels. The previous chapter showed how each variable listed in Table 4-1 could be manipulated to indicate four levels of alignment within OLMs. Patterns have been used from Table 3-2, and applied to the skill meter visualisation learner-model as seen in Figure 4-1. Different levels of misalignment for each visual variable is indicated on three skill meters with each skill meter has three segments (known, misconception and un-known). This is done to enables students to differentiate the level of misalignment in each segment. The skill meter is based on the OLMlets OLM that has been used in the context of this thesis to fulfil the studies aims.

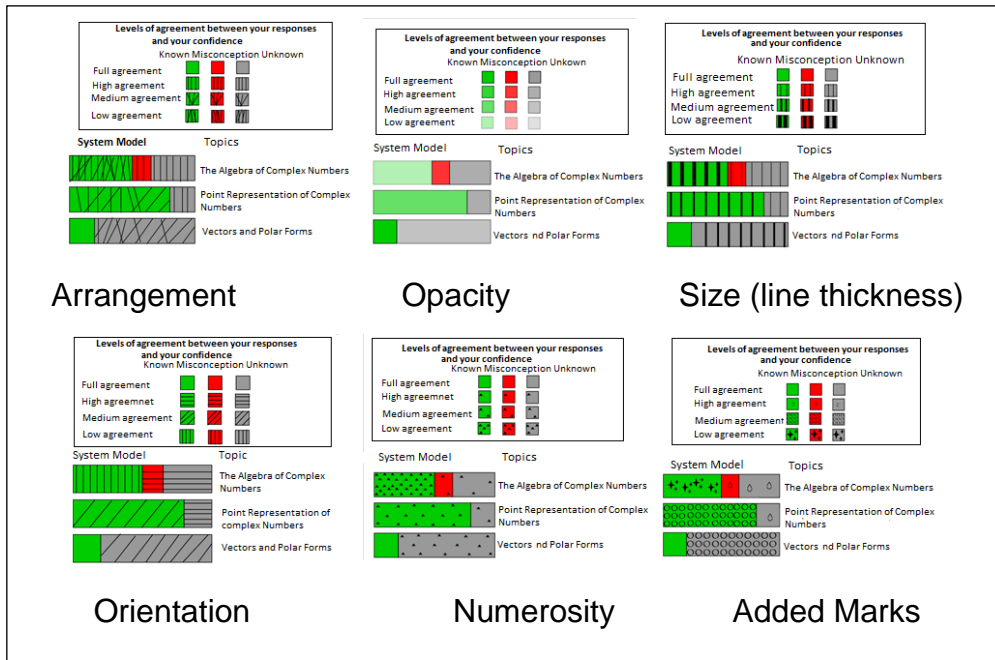


Figure 4-1. Visualising alignment using six different methods of visual variables applied to a skill meter learner model.

Figure 4-1 shows the figures provided to students in this exploratory study explained in the next section, to indicate the preferred method for visualising alignment. The word ‘agreement’ was used instead of ‘alignment’, so students can easily interpret the information and understand the levels of agreement shown by the different levels, depicted by the visual variable. All the figures include the same information, but each is indicated with a visual variable to indicate the level of alignment. The visual variable ‘arrangement’ shows that the first topic (The algebra of complex numbers) has untidy lines in the green segment of the skill meter, which indicates a higher misalignment level, representing that the student was unconfident in their knowledge. The more organised the lines, the more aligned the students’ perception to the system evaluation about their knowledge.

The second visual variable is the opacity of the colour. The higher the transparency, the higher the misalignment. It can be seen that the first topic is highly transparent in the green segment of the skill meter. The third variable is the size of the line (thickness), the thicker the line, the higher the misalignment. Orientation of the lines represents the level of misalignment, where the nearer to vertical the lines, the higher the misalignment. The visual variable numerosity is represented by the dots in each segment, where the greater the number of dots shown in each segment indicates a higher level of misalignment. The sixth visual variable is added marks. In this visual variable, each level of misalignment is represented with a special mark that shows that it belonged to the same level of misalignment.

Based on Gestalt principles, people can perceive each visual variable differently (Ware 2004). Therefore, it is important to explore the preferred method and consider students' perspectives on the visual variables used to indicate alignment.

## **4.5 Method**

This study aims to explore students' perceptions of their preferred method from the six visual variables discussed earlier, and presented in Figure 4-1, to visualise alignment in the context of OLMs. Visualising alignment shows how student self-confidence is aligned with a computer evaluation of the student's performance.

### **4.5.1 Study design**

The study utilised a questionnaire including 15 paired comparison questionnaire items. Figure 4-1 shows the six visual variables applied in skill meter. The

questionnaire also included questions about students' experience of using OLM, with additional questions about their gender and level of study (undergraduate or MSc).

The visual variables that appeared in the paired comparison questionnaire items were randomly ordered, to ensure that the order of the visual variables could not affect students' selection.

#### **4.5.2 Participants**

Participants were students recruited from the School of Engineering at the University of Birmingham, UK ( $N = 67$ ). There was ( $n = 42$ ) undergraduate students and ( $n = 22$ ) master students. There were 45 male students and 18 female students from the 67 students who participated in the study. There were 41 students who experienced using the OLM and 22 had no experience in OLM. There was some missing information about the base questions of gender, OLM experience and study level.

#### **4.5.3 Study procedures**

Questionnaires were printed in colour (as shown in Appendix 1) and distributed to participants. Whilst the figures in the questionnaires were rather small, I did not receive any complaints from students about their size. A 15-minute presentation was given to participants, to help them to understand the visualisation and how each visual variable was used to indicate the level of alignment. This was to ensure that students could clearly interpret the visual variables so that their selection

would be based on their preferred method, not their understanding (although these two terms could not be completely separated).

Each participant had 15 pairs of the skill meter visualisations and each participant was required to select one item from each pair. Completed surveys where a participant failed to complete all of the 15 paired comparisons was excluded from the analysis. As this study used a forced-choice task, the full survey should be completed.

#### **4.5.4 Analysis**

Two methods were used to analyse the survey data. The first method used Thurstone's paired comparison method (Thurstone, 1927). This method uses binary choice to measure students' preference of items (i.e., visual variables), by giving all possible pairs of items to each respondent who choose their preferred choice from each pair. Each respondent provides a preference score for each item. The scores are calculated by creating a  $t$  by  $t$  matrix, and a score of 1 is entered in each cell where the column item was preferred to the row item, or 0 otherwise (for a review, see Thomas et al., 2009). The Thurstone paired comparison method (Thurstone, 1927) uses a scale of preference to show the most and least preferred methods.

The second method used was the statistical method, which is used in all studies presented in this thesis. The Shapiro-Wilk test (Shapiro and Wilk, 1965) was used to determine the normality of the data. For normally distributed data, T-tests are used between two groups (male, female) or between two sets (pre-test, post-test)

to compare the mean score on a continuous variable. There are two types of t-tests: the independent t-test, which is used to compare between two different independent groups, and the paired sample t-test, which is used to identify a change in a score from Time 1 and then repeated again at Time 2. If the data is not normally distributed, the Mann-Whitney U-test is equivalent to an independent t-test and the Wilcoxon signed rank test is equivalent to a paired sample t-test. In parametric tests, the mean and standard deviation are used, whereas non-parametric test uses the mean rank. The ranking of the data gives the lowest score (a score of 1) and then finds the next highest score and give it a score of 2, and so on. The mean rank shows the average score of each rank assigned for each item. Then the ranks are used to identify the significant difference between the two groups.

To identify the significant difference for two or more groups, a one-way ANOVA is used for normally distributed data to compare the mean score between the groups. The one-way ANOVA shows if there is a significant difference between the groups but it will not specify which groups actually differ. Thus, a post-hoc analysis such as a t-test should be conducted, to identify the difference between the two groups. There are two types of one-way ANOVA, as follows: a repeated measure ANOVA where the same participant is compared on more than two occasions; the second type is a between-groups ANOVA, used when comparing the mean of two or more different participants. The equivalent test for the one-way between groups is the Kruskal-Wallis test, and the equivalent test for the one-way repeated measure is the Friedman test.

#### **4.5.5 Data collection**

For the paired comparison method, the data was entered into a Word Excel spreadsheet, where a 6 × 6 cell matrix was created. All scores from the survey were entered into the matrix.

For the statistical method, data from the survey were entered into SPSS (a statistical data analysis program) based on their preferred score from the paired comparison questionnaire, see Appendix 2. There was one independent variable (student ID) and six dependent variables (the six visual variables used in this study). Each participant had 15 selections of their preferred method, which were distributed based on their preferred selection from each pair. The total number of times the participant preferred the visual variable over the other variables was entered under the preferred dependent variable.

#### **4.6 Results**

Students' preferred method for visualising alignment was analysed using the Thurston's paired comparison method and inferential statistical analysis. The results from each method are described in detail in the following sub-sections.

##### **4.6.1 Thurstone's paired comparison method**

Six visual variables were ordered in the matrix as 6 × 6. Each participant assigned a score of 1 for their preferred method (of the two visualisations in each paired comparison item) under the preferred column linked with the visual variable to the row.

In Table 4-2, the results are presented for the preferred method of visual alignment. This shows that there were 61 participants who preferred the visual variable opacity to the visual variable arrangement, whereas only 6 participants preferred the visual variable arrangement to opacity. The total score of the preferred visual variable over all other variables is indicated by the sum for each preferred column, linked with each visual variable, as shown in Table 4-2.

*Table 4-2. Applying Thurstone's paired comparison method for participants' preferred method of visualising alignment.*

	<i>Arrangement</i>	<i>Opacity</i>	<i>Size</i>	<i>Orientation</i>	<i>Numerosity</i>	<i>Added Marks</i>	<i>Sum</i>
<i>Arrangement</i>		61	52	60	53	54	280
<i>Opacity</i>	6		17	13	17	21	74
<i>Size</i>	15	50		36	39	36	176
<i>Orientation</i>	7	54	31		38	29	159
<i>Numerosity</i>	14	48	29	29		33	153
<i>Added Marks</i>	13	47	32	38	33		163
<i>sum</i>	55	260	161	176	180	173	1005

The sum score indicated under the preferred column shows the total score of the preferred method over all other visual variables. For example, the opacity of the colour had the highest preferred score of 260 over all other visual variables whereas arrangement had the lowest preferred score of 55.



There was a total number of responses of 1005 for all participants (67 participants × 15 paired comparison items) and each score for each visual variable was divided by the total response, as shown in Table 4-3.

*Table 4-3. Total number of responses for each visual variable divided by total responses.*

	<i>Arrangement</i>	<i>Opacity</i>	<i>Size</i>	<i>Orientation</i>	<i>Numerosity</i>	<i>Added Marks</i>
<i>Arrangement</i>	0.00	0.06	0.05	0.06	0.05	0.05
<i>Opacity</i>	0.01	0.00	0.02	0.01	0.02	0.02
<i>Size</i>	0.01	0.05	0.00	0.04	0.04	0.04
<i>Orientation</i>	0.01	0.05	0.03	0.00	0.04	0.03
<i>Numerosity</i>	0.01	0.05	0.03	0.03	0.00	0.03
<i>Added Marks</i>	0.01	0.05	0.03	0.04	0.03	0.00

Data were then normalised using Z-transform, as shown in Table 4-4. Using z-score makes the comparison between variables easier. The z-score indicates exactly where the score is located relative to all other scores in the distribution. The smaller the sum of the z-score for each visual variable, the more preferred the method, whilst the larger the sum of the z-score, the least preferred the method. The preference score for each visual variable is indicated by the sum of each z-score for each column linked with each visual variable, as shown in Table 4-4.

When applying the Thurstone paired comparison method (Thurstone, 1927), a scale of preference is used where each visual variable is ordered on the scale from least to most preferred.

Table 4-4. Normalising data using Z-transform for each visual variable.

	Arrangement	Opacity	Size	Orientation	Numerosity	Added Marks
Arrangement		0.81	1.45	1.47	1.23	1.41
Opacity	0.54		0.57	0.78	0.69	0.44
Size	1.00	0.31		0.32	0.48	0.40
Orientation	0.37	0.49	0.24		0.43	0.01
Numerosity	0.83	0.21	0.13	0.02		0.23
Added Marks	0.66	0.17	0.30	0.41	0.16	
Sum	3.39	1.98	2.69	3.00	2.99	2.49

Figure 4-2 shows a preference scale of the visual variables, based on sum of the z-score.

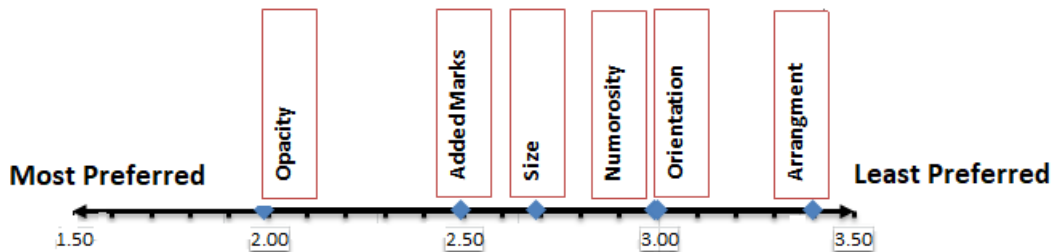


Figure 4-2. Scale of students' preferred method from most to least preferred.

Opacity was the most preferred method, indicated by the lowest z-score (1.98). The visual variables (Added Marks = 2.49) and (Size = 2.69), were closer together in the scale of preference, both of which were more preferred than Numerosity (2.99) and Orientation (3.00), which almost scored the same. (Thus, the blue dots on the scale for both variables were very close, which appeared as one big dot). Arrangement was the least preferred method of visual variable, which indicates that participants did not prefer the inconsistent lines.

#### **4.6.2 Statistical analysis**

The second method used was statistical analysis. This method was used for stronger argument as it was also used in all studies presented in this thesis, and is widely used in research. Descriptive statistics were calculated for each visual variable, as shown in Table 4-5.

Table 4-5. Descriptive statistics for the preference for each visual variable.

	N	Mean	SD	Percentiles		
				25th	<i>Mdn</i>	75th
Arrangement	67	.82	1.058	.00	.00	1.00
Opacity	67	3.88	1.387	3.00	5.00	5.00
Size	67	2.40	1.280	1.00	2.00	3.00
Orientation	67	2.64	1.252	2.00	3.00	4.00
Numerosity	67	2.70	1.393	1.00	3.00	4.00
Added Marks	67	2.55	1.470	1.00	3.00	4.00

The Shapiro-Wilk test was used to determine the normality of the data. Data were not normally distributed ( $p < .05$ ). The Friedman test is non-parametric test used to detect differences of items among multiple items. Thus, it was used to perceive if there was a significant difference between the six visual variables. By applying the Friedman test, a significant difference between the six variables (based on participants' preferred method) was observed ( $\chi^2 (5) = 104, p < .001$ ) with a mean rank of (Arrangement = 1.76), (Opacity = 4.96), (Size = 3.42), (Orientation = 3.67), (Numerosity = 3.61) and (Added marks = 3.58). The ranking shows the variables that had the lowest scores were the least preferred and those with the highest scores were the most preferred. Thus, the Arrangement variable had the lowest mean rank score and Opacity had the highest mean rank score. This result corroborates the result carried in the previous test.

A post hoc analysis should be conducted to identify any difference in preference within pairs of visual variables; therefore, the Wilcoxon signed rank test was used. This is a non-parametric test used to compare two items. Table 4-6 shows a 6 x

6 matrix for the comparison between each pair of visual variables, based on participant preference, using the Wilcoxon signed-rank, indicating whether there is a significant difference between the two variables by (\*) with the z-value.

*Table 4-6. A 6 × 6 matrix showing the z-value of Wilcoxon signed-rank assigned to the top row for the higher mean rank for the associated visual variable.*

	<b>Arrangement</b>	<b>Opacity</b>	<b>Size</b>	<b>Orientation</b>	<b>Numerosity</b>	<b>Added marks</b>
<b>Arrangement</b>		-6.66 *	-5.69 *	-6.15 *	-5.8 *	-5.14*
<b>Opacity</b>						
<b>Size</b>		-4.81 *				
<b>Orientation</b>		-4.13 *	1.23		-.175	-.37
<b>Numerosity</b>		-4.08 *	-1.23			.59
<b>Added marks</b>		-3.95 *	-.47			

\*represents a significant value  $p < .001$ .

Comparing each pair of visual variables, Opacity was significantly preferred among all other visual variables using Wilcoxon signed-rank test, and Arrangement was statistically least preferred than all other visual variables. The result indicates that participants tended to prefer the intrinsic technique (Opacity). It also reflects that students did not prefer the unorganised lines (Arrangement) to other organised lines (e.g. Size and Orientation).

## 4.7 Discussion

The main aim of this study was to explore students' preference method of visual variables to indicate the alignment within OLMs. This survey-based study suggested that students preferred Opacity as a method to visualise alignment. Opacity has been considered an intrinsic technique, whereby the visual variable is used by altering the existing display. Opacity was preferred among all other visual variables such as Size, Added Marks, Numerosity, Orientation, and Arrangement, which used extrinsic techniques to indicate the level of alignment. This study is considered a novel study for exploring students preferred method to visualise alignment within OLMs. The results are in line with other research from the field outside of the educational domain, with regard to the popularity of opacity as a means to visualise uncertain data (Kinkeldey, MacEachren and Schiewe, 2014). Similarly, users in fields other than education preferred well-structured visualisation that indicate uncertainty information over the less well-organised visualisation (Boukhelifa et al., 2012), which is in line with the result of this study, where Arrangement was a less preferred option among all other visual variables.

This study answers the second research question of this thesis concerning the students' preferred method to visualise alignment for OLM users. Knowing students' preferred method in visualising alignment could provide ideas about students' preferences, which could motivate students' use within OLMs. Researchers argue that participants can better retrieve the information (illustrated by the participants preferred method) to indicate variability or inconsistency in the data. The findings of this study can enable investigation of the effect of the alignment

using a control condition on students' knowledge monitoring skills when students are practically using their preferred method within OLMs. This is discussed in the evaluative studies in Chapters 5 and 6. In addition, Chapter 7 compares the effect of the alignment within OLMs on students' knowledge monitoring skills when comparing visually presented OLMs using the students' preferred method and textual description of the OLM.

# **Chapter 5 : EVALUATING THE EFFECT OF VISUALISING ALIGNMENT ON STUDENTS' KNOWLEDGE MONITORING SKILLS**

---

## **5.1 Overview**

This chapter presents an empirical study that aims to explore the impact of visualising alignment between students' confidence in the correctness of the answers and their performance on answers given. This is based on students' confidence in the pre- and post-test, and their behaviour in using the system, such as the number of questions answered and the number of OLM views.

A pre-questionnaire was used to identify students' perceptions on items related to learning and confidence before using the OLMlets system. A post-questionnaire was then used to explore students' perceptions on items related to learning, confidence, and their perception of using the system after OLMlets use. In addition, pre- and post-tests were used to explore students' knowledge before and after OLMlets use. Students were randomly divided into three groups: two treatment condition groups and a control condition group. The two treatment condition groups were used to explore the impact of the presentation of the alignment on students' confidence and on their engagement in using the



system. This study was conducted in real class environment, to answer one of the key research questions of this thesis: how does visualising alignment support students' confidence judgement?

## **5.2 Visualising alignment in OLM**

Knowledge monitoring skill is a fundamental component of metacognition. Researchers have indicated the significance of knowledge monitoring skill and its relation to academic achievement (Jackson and Kleitman, 2014). However, there are few investigations on the influence of the visualisation when the OLM incorporates student confidence in their knowledge and the system assessment of their knowledge.

For example, in Mr.Collins (Bull and Pain, 1995) and CALMsystem (Kerly, Ellis and Bull, 2008), students are allowed to compare their own rating of their understanding of a topic with those inferred by the system. These systems developed OLMs that encouraged negotiation to facilitate students' reflection on their ability to monitor and evaluate their understanding, and to improve the accuracy of the system's model. Negotiation is employed in intelligent tutoring systems when a conflict between the ratings of the system and the learner occurs (Lan et al., 2008). During such negotiation, the system seeks to achieve a balance between its rating of learner activity (i.e., the system's beliefs concerning learner understanding and ability) and the rating provided by the learner. The negotiation follows the idea of interaction symmetry where the system and the student are able to make negotiation moves. For example, both the system and the learner can initiate, maintain and end negotiation. Often the negotiation dialogue will allow

each party (system or learner) to provide some justification of, or explanation for, their viewpoint, or to challenge or amend the beliefs expressed by the other party. Negotiation can either party to accept, compromise or maintain beliefs if they did not reach agreement (Bull and Pain, 1995; Kerly, Ellis and Bull, 2008). One way of ending negotiation is when the learner can defend their beliefs by answering further test questions or by accepting the system justification. Then the system updates the learner model with the outcome of the negotiation (Suleman et al. 2015). Figure 5-1 shows a transcript of a typical conversation fragment based on CALMsystem (Kerly, Ellis and Bull, 2008) where a conflict between the system and the learner occurred in a topic called Friction.

User:

now friction

OLMBot:

I believe that you have a high knowledge level for the Friction topic. You have said that you have a moderate confidence level in your ability for this topic. We still need to resolve this difference. Would you like to:

1: change your beliefs so that you agree with me (The recommendation is high knowledge level) OR

2: see why I hold my views (have me explain) OR

3: view your and my beliefs about your knowledge OR

4: see if we can agree halfway OR

5: answer some questions to show me how much you know?

User:

2

OLMBot:

You answered 3 question(s) and got 0 wrong for the Friction topic. You need to choose from one of the following options. Would you like to :

1: change your beliefs so that you agree with me (The recommendation is high knowledge level) OR

2: view your and my beliefs about your knowledge OR

3: see if we can agree halfway OR

4: answer some questions to show me how much you know?

User:

Answer questions

OLMBot:

OK, I'll ask some test questions for you to demonstrate your knowledge. Here's the first

...(ctd).

*Figure 5-1. CALMsystem OLM (Kerly, Ellis and Bull, 2008) conversation fragment between the system and the learner to resolve a conflict occurred in a topic called Friction using chatbot.*

Although, Mr.Collins, CALMsystem and VisMod (Bull and Pain, 1995; Kerly and Bull, 2008; Zapata-Rivera and Greer, 2004) used a negotiation mechanism to resolve a misalignment, the system used in this study uses a simple inspectable OLM.

Moreover, the study presented in this chapter uses a control condition to test the impact of visualising alignment in the OLM, where other studies did not use the process of a control condition study.

Based on the theory discussed in the Chapter 1, providing students with feedback about their confidence and their performance might be more effective than only giving the students feedback about their score of assessment (Butler and Winne, 1995). Further, visualising alignment may influence students' behaviour in using the system, as they are given the opportunity to observe misalignments between their performance and their confidence when solving an assessment. As a result, students could be more motivated to resolve the misalignment (Kulhavy and Stock, 1989).

Accordingly, the hypotheses for this study are as follows:

**H1:** Visualising alignment in the OLM will improve students' confidence.

**H2:** Visualising alignment in the OLM will motivate students to answer more questions and to view their learner model more.

The study explores students' confidence related to their use of OLMlets and their interaction using the system, while observing the alignment information using a control study.

## 5.3 Method

The method section illustrates the different versions of the OLMlets system (discussed in Chapter 3) that were implemented in the system for the purpose of this study and how information of the alignment is different in each version. Study design, procedures, analysis, instruments, and participants are described.

### 5.3.1 The OLMlets OLM

In this study, three versions of the learner model representation have been developed in OLMlets as described in Chapter 3. The study has three condition groups (two treatment condition groups and a control condition group). The control condition group uses the standard skill meter used in OLMlets where the alignment information is hidden from the learner (Figure 5-2). Both treatment condition groups show the alignment between the computer's evaluation and how confident the students are in their answers that the answers are correct.

**Control condition (baseline condition):** The control condition uses the standard skill meter used in the OLMlets OLM, where only the system's evaluation of the students' performance on the system's automated assessment is shown (Figure 5-2). However, students are still requested to assess their confidence after each question.

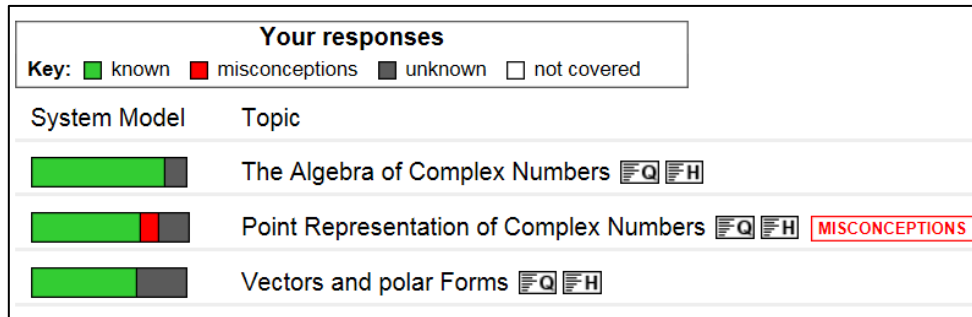


Figure 5-2. Skill meter visualisation (Control condition).

**Treatment condition 1 (The combined model):** The first treatment condition is the combined model. This condition uses the opacity of the colour to indicate the level of misalignment between students' confidence and the computer evaluation about the student performance. The fill colour of the segment in the skill meter shows high transparency when there is high misalignment (low agreement) and the more opaque the colour for more aligned (high agreement). In Figure 5-3, the first and third topics show the green colour is fully opaque, which indicates that students' confidence in their performance is highly aligned with the computer evaluation about the student performance. In the second topic (Point Representation of Complex Numbers) a misalignment is indicated by the transparency of the green colour, which specifies that the value of model alignment has a score less than 0.3 (low agreement) for the known concept (as discussed previously in Chapter 3). The grey colour indicates that students experienced some difficulties in answering the system's automated assessment. The grey colour in the first and second topics indicates a medium level of agreement (medium misalignment). This reflects that the student believes that they know the answers when

they performed incorrectly. In contrast, the third topic shows a more opaque grey colour (Vectors and Polar Forms), which shows that the student accurately judged their confidence based on their performance, which is aligned with the computer evaluation about their performance. Whilst the red colour in topic 2 indicates that the student holds a misconception, the transparent red colour indicates that the student was not sure about the answer, which means that the answer might have been a guess, rather than a misunderstanding of the concept.

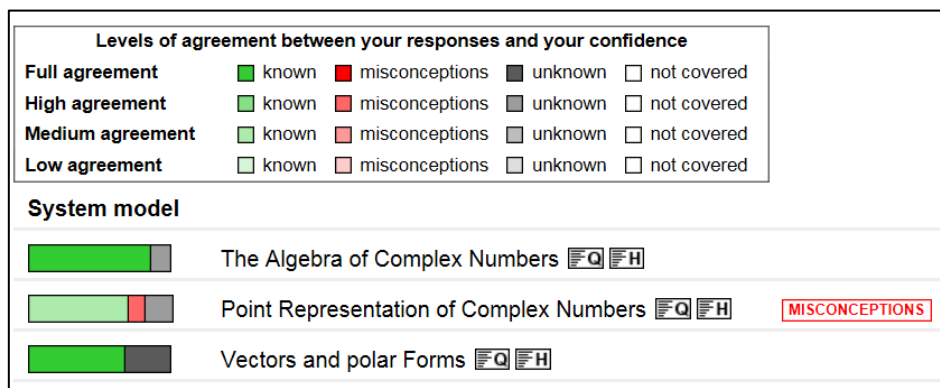


Figure 5-3. The combined model skill meter uses the opacity of the colour to visualise alignment.

**Treatment condition 2 (The expandable model):** This model uses the combined model as the base visualisation with an additional feature that allows the students to expand the model to two separated skill meters. In this condition, students can compare directly between the two models (your knowledge, your confidence) by clicking on the button 'show models'. If the student clicked the show models (as seen in Figure 5-4), the model is expanded (Figure 5-5).

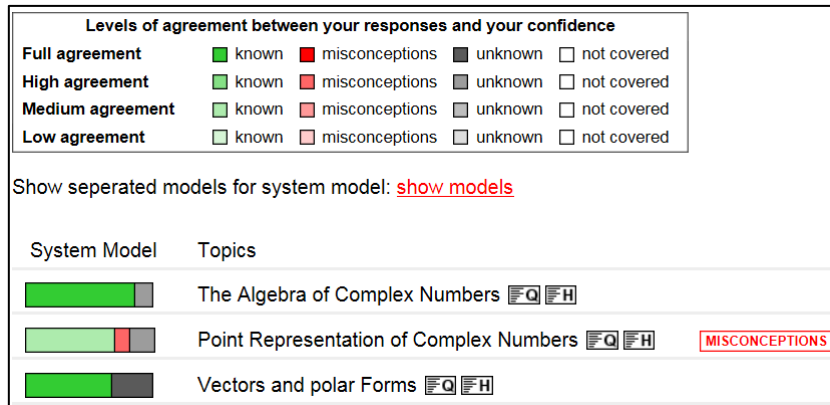


Figure 5-4. The expandable model when the model is combined.

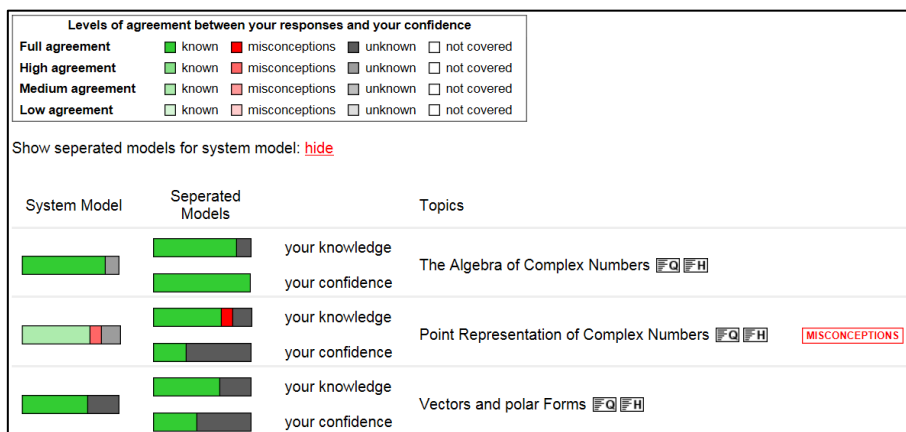


Figure 5-5. The expandable model when the model is expanded.

### 5.3.2 Study design

To study the impact of visualising alignment in a real class setting, the domain in OLMlets was based on an undergraduate course related to mathematics (*Introduction to Complex Variables*). The study took place over a three-week period at the beginning of the academic year, and students used the system to prepare for their first class test. Three topics were integrated into OLMlets from the course



syllabus. Each topic included 25 questions. Students were randomly assigned to the three conditions discussed earlier. The control condition group used the baseline skill meter (Figure 5-2). The two treatment condition groups used the combined model condition (Figure 5-3), or the expandable model condition (Figure 5-5).

### **5.3.3 Study procedures**

As the study was based on a mathematics course (*Introduction to Complex Variables*), the course instructors agreed to comply with the study procedures and re-check the questions integrated into the OLMlets with their answers. A presentation of approximately 15 minutes was given (from class time) to explain the system use and its features, and to explore how the alignment between the computer evaluation about their performance and the students' confidence was visualised. Students were given a username and password to login to the system via the website <http://OLMlets.bham.ac.uk>. The system was accessible at all times.

A user manual for the three versions of the OLMlets system—with step by step printed screenshots for each version of the OLMlets—were distributed based on which versions the student belonged to. Students used the system voluntarily in their own time as a preparation tool for their class test.

At the start of the study, a pre-questionnaire was circulated to explore students' perceptions about their metacognitive skills before using the system. After students used the system, a post-questionnaire was distributed to explore students' metacognitive skills after using the system. Further, a pre- and post-paper-based

test (that included the same strategy as the OLMlets assessment) was used to measure students' confidence judgement before and after using the system.

#### **5.3.4 Instruments**

The pre-questionnaire consisted of two sections (see Appendix 3). The first section included items related to learning, such as how to plan and monitor ones' own learning. The second section included items that explored students' perceptions about their confidence while learning. The post-questionnaire consisted of four section. The first two sections were the same for all conditions, focusing on two aspects: planning and monitoring learning; and students' confidence (see Appendix 5). A section including items about visualising alignment in the OLM was included in the third section. This section was only answered by students of both treatment conditions (combined and expandable model), see Appendix 6. The last section was for students from the expandable model condition, where the questionnaire's items were related to how the model was expanded, and whether seeing the two separated models affected their behaviour in using the system and their knowledge monitoring skills (see Appendix 7). A paper-based pre- and post-test were provided to students, with the same strategy of assessing the students' confidence after answering each question as used in the OLMlets (see Appendices 9 and 10). The study was with students in a real class situation, and as such, the time set aside in class for the researcher was limited. The study did not check if questions for both pre- and post-tests were at the same level of difficulty (students' confidence judgement might be influenced if one test was more difficult than the other).

Owing to the small number of participants in the study and the limited time available in the class, the study did not use a standardised test. These are used for large numbers of students and use special answer sheets that can be read by computer to compare between students regarding students' achievement. However, all students had the same pre- and post-test, which contained 10 multiple-choice questions. All tests were graded in the same way, with a maximum score out of ten, and any incomplete answer being marked as incorrect, as this is the procedure followed in a real class setting.

### **5.3.5 Study analysis**

A Shapiro-Wilk test was conducted to assess the normality of the data, where ( $p > .05$ ) is considered a normal distribution. Students' performance in the pre- and post-tests followed the same criteria as in the actual class test, where unanswered questions are considered incorrect. The questionnaire data were analysed using Kruskal-Wallis tests to explore any significant differences between conditions. Non-parametric tests use mean rank instead of the average (usually used in parametric tests). Ranking the data gives the lowest score a score of 1 and then find the next highest score and give it a score of 2, and so on. The mean rank shows the average score of each rank assigned for each item. Then, the ranks are used to identify the significant difference between the two groups.

Students' confidence gain was calculated, based on the difference between the self-confidence assessment from the post- and pre-tests to investigate the relative impact of the three visualisation conditions within OLMlets on students' confidence. To identify the relationship between students' interaction with the system

(the number of OLM views and number of questions answered) and their pre- and post-confidence, Pearson's correlations were used.

### **5.3.6 Participants**

Participants for this study were undergraduate students ( $N = 110$ ) from the Sultan Qaboos University in Oman. There were 110 students registered on the course. Students volunteered to use the system, and a total of 79 decided to participate, signing the informed consent participation form. Students were assigned randomly to each of the three conditions: control condition ( $n = 27$ ), treatment condition A (combined model,  $n = 27$ ), and treatment condition B (expandable model,  $n = 25$ ). In total, 54 participants completed both the pre- and post-tests; however, because participation was voluntary, only 38 actually logged on to OLMlets. Furthermore, not all participants had valid logins; subsequently, those who did not view the model ( $n = 13$ ) were excluded from the conditions of OLMlets. This is because participants who did not view the model could not benefit from accessing their learner model to promote their metacognitive skills. This suggests that students were only answering a formal assessment of multiple-choice questions without receiving feedback about their level of knowledge, so they were not considered as OLMlets users. Table 5-1 shows the number of participants belonging to the two conditions: those who used the OLMlets and viewed the model ( $n = 25$ ) and those who did not use the OLMlets or did not view the model ( $n = 29$ ). The number who participated in each condition are shown in Table 5-1.

Table 5-1. Number of students participating in each condition

	No-OLMlets	OLMlets (N = 25)		
		Control	Combined	Expandable
<b>Number of students</b>	29	9	9	7

## 5.4 Results

The study first explored the benefit of using the OLMlets system on students' knowledge monitoring skills and performance by comparing non OLMlets users to students who used the OLMlets system. The study then focused on the three visualisation conditions and their impact on student confidence and their interaction with the system. Appendix 11, shows students logs data for all students for the data used in this study.

### 5.4.1 Normality of data

The Shapiro-Wilk test showed that assessing the normality of data between the two groups of OLMlets-users and no-OLMlets users determined that pre- and post-confidence were normally distributed ( $p > .05$ ), thus a parametric test was used. Students' pre-test data was not normally distributed; however, the post-test was normally distributed. Student confidence gain was not normally distributed for the group of OLMlets-users, thus a non-parametric test was used ( $p < .05$ ).

Among the sub-groups of OLMlets users (which included the three visualisation conditions), the Shapiro-Wilk test indicated that pre-confidence was normally

distributed ( $p > .05$ ), but post-confidence was not normally distributed among the conditions ( $p < .05$ ), thus a non-parametric test was used. Participants' confidence gain from their pre-confidence to their post-confidence, pre-test score and participants' post-test score were normally distributed ( $p > .05$ ); thus, a parametric test was used. Participants' interaction with the system in terms of number of questions answered and number of times the model was viewed was not normally distributed ( $p < .05$ ); thus, a non-parametric test was used to identify significant differences between the groups.

#### **5.4.2 Student awareness from the pre- and post-questionnaire**

Participants' perceptions of their own metacognitive skills were explored using the questionnaire data. Questionnaires were distributed before and after using the OLMlets system, to identify the differences that may occur as a result of using the system in terms of learning or confidence in answering the system's automated assessments. The pre-questionnaire (see Appendix 3) was divided into two sections: perceptions of learning before OLMlets use, and perceptions of confidence before OLMlets use. Participants' responses for the pre-questionnaire for all groups were entered into SPSS file (a statistical data analysis program), see Appendix 4. Participants' responses for the post-questionnaires for all groups were also entered into SPSS file, see Appendix 8.

**Students' perception on learning before OLMlets use**

Table 5-2 presents the results from the first section of the pre-questionnaire related to learning, and the inferential statistics using the Kruskal-Wallis test to identify the significant difference between groups and the mean rank for each group. The higher the mean rank shows that participants had a greater agreement with the item.

*Table 5-2. Shows the mean rank for questionnaire items based on the difference between the three groups and the result of the Kruskal-Wallis test. The significant items are identified using bold text.*

	<b>Baseline</b>	<b>Combined</b>	<b>Expandable</b>	$X^2(2)$	<i>P</i>
Doing homework helps me see how much I know.	21.78	19.54	29.00	5.00	0.082
<b>Taking tests helps me to identify gaps in my knowledge.</b>	<b>23.78</b>	<b>16.38</b>	<b>29.33</b>	<b>8.39</b>	<b>0.015</b>
<b>Taking tests helps me to identify my misconceptions.</b>	<b>23.44</b>	<b>16.04</b>	<b>27.10</b>	<b>6.37</b>	<b>0.041</b>
Doing homework helps me to monitor my learning.	23.36	23.88	23.33	0.018	0.991
I use tests and homework to plan for my learning.	20.92	20.58	27.93	3.23	0.199
I let the teacher or my parents plan for my learning.	20.86	24.35	25.93	1.34	0.511
I plan for my own learning.	26.56	19.62	23.20	2.42	0.298

From the results in Table 5-2, there were two questionnaire items that had significantly different views. The first item was ‘*Taking tests helps me to identify gaps in my knowledge*’, where participants stated different views ( $\chi^2(2) = 8.39, p = .015$ ). The expandable model condition had the highest mean rank of 29.33, where the baseline condition had a mean rank of 23.78. The combined model had the lowest mean rank of 16.38. The second significant item was ‘*Taking tests helps me to identify my misconceptions*’, using Kruskal-Wallis test ( $\chi^2(2) = 6.37, p = .041$ ). The expandable model and baseline conditions had higher mean ranks of 27.10 and 23.44, respectively, compared to the combined model condition with a mean rank of 16.04. Based on these results of the pre-questionnaire, participants from the combined model condition were more likely to express that tests did not support them to identify their lack of knowledge or misconception. Figure 5-6 shows the percentage distribution of participants’ perceptions for items that found significantly different views between conditions in the learning section of the pre-questionnaire based on the 5-point Likert scale.

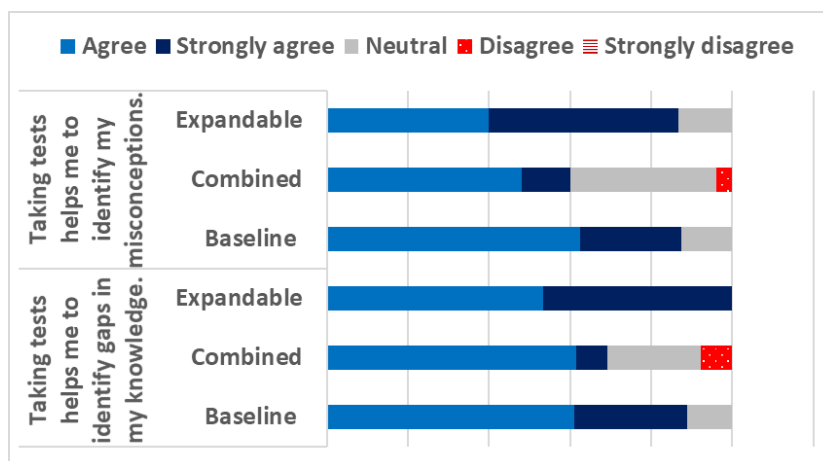


Figure 5-6. Pre-questionnaire items where different views were found from the learning section before OLMlets use.



**Students' perception on confidence before OLMlets use**

Table 5-3 shows the second section of the pre-questionnaire related to confidence and the inferential statistics using the Kruskal-Wallis test to identify the significant difference between groups, and the mean rank for each group.

*Table 5-3. The mean rank for each group on the pre-questionnaire items related to their confidence and the significant items is identified using the bold text.*

	<b>Baseline</b>	<b>Combined</b>	<b>Expandable</b>	$\chi^2(2)$	<i>P</i>
I am good at assessing my confidence level.	21.44	22.08	27.20	2.13	0.345
I try to increase my knowledge when my confidence is low.	19.94	20.33	28.80	5.53	0.063
<b>I try to increase my knowledge when my confidence is high.</b>	<b>22.85</b>	<b>16.19</b>	<b>29.07</b>	<b>7.57</b>	<b>0.023</b>
My confidence increases when I do well on a test or homework assignment.	24.08	18.88	26.80	3.37	0.186
I feel my confidence decreases when I do poorly on a test.	23.11	23.08	24.33	0.10	0.954
I know how well I will do before I write a test.	20.61	24.35	26.23	1.64	0.441

There was only one item on the questionnaire that revealed a significant difference, 'I try to increase my knowledge when my confidence is high',  $\chi^2(2) = 7.57$ ,  $p = .023$ , with mean ranks of 29.07, 22.85, and 16.19 for the expandable model, the

baseline, and the combined model conditions, respectively. Participants from the expandable model and the baseline conditions seemed to learn more when their confidence was high, compared to students from the combined model condition who felt that high confidence was not influenced by their willingness to increase their knowledge. Figure 5-7 shows the percentage distribution of participants' perceptions in the confidence section of the pre-questionnaire based on the 5-point Likert scale for the pre-questionnaire item that revealed a significant difference related to confidence.

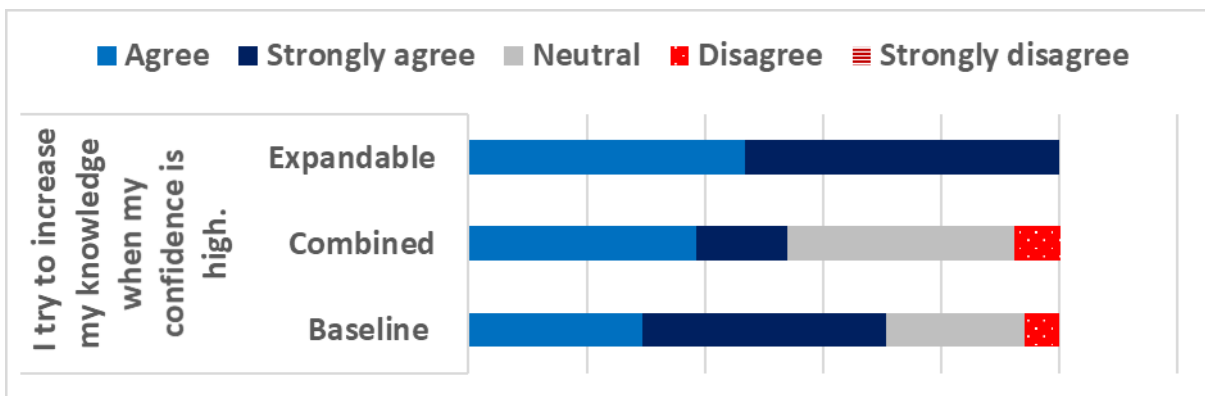


Figure 5-7. Pre-questionnaire item for confidence section where different views among students was found before OLMlets use.

### **Students' perception on learning after OLMlets use**

Table 5-4 shows the questionnaire items from the post-questionnaire related to learning after OLMlets use with the inferential statistics using the Kruskal-Wallis test and the mean rank used in non-parametric tests.

Table 5-4. The mean rank for each group on the post-questionnaire items related to their learning.

	Baseline	Combined	Expanded	$\chi^2(2)$	<i>P</i>
OLMlets helped me to see how much I know.	10.67	11.71	12.50	0.41	0.815
OLMlets helped me to identify my gaps in my knowledge.	9.28	12.93	13.17	2.32	0.314
OLMlets helped me to identify my misconceptions.	9.94	10.43	15.08	3.03	0.220
OLMlets helped me to monitor my learning.	10.61	10.79	13.67	1.13	0.567
OLMlets helped me to plan for my learning.	10.22	11.50	11.67	0.33	0.846
<b>OLMlets encouraged me to answer more questions.</b>	<b>7.39</b>	<b>14.64</b>	<b>14.00</b>	<b>7.19</b>	<b>0.027</b>

Figure 5-8 shows the percentage distribution for students' preception about the questionnaire item '*OLMlets encouraged me to answer more questions*'. This questionnaire item revealed a significant difference between the three conditions using Kruska-Wallis test ( $\chi^2(2) = 7.19, p = .027$ ), with mean ranks of 7.39 for the baseline condition, 14.64 for the combined model condition, and 14.00 for the expandable model condition. Both treatment conditions (who could visualise alignment in their OLM) had similar beliefs that OLMlets encouraged them to answer more questions, compared to the control condition, who had different views. This result may indicate that the baseline condition (where participants could not view the

alignment between their confidence and the computer evaluation about their performance) was less motivating for participant interaction with the system.

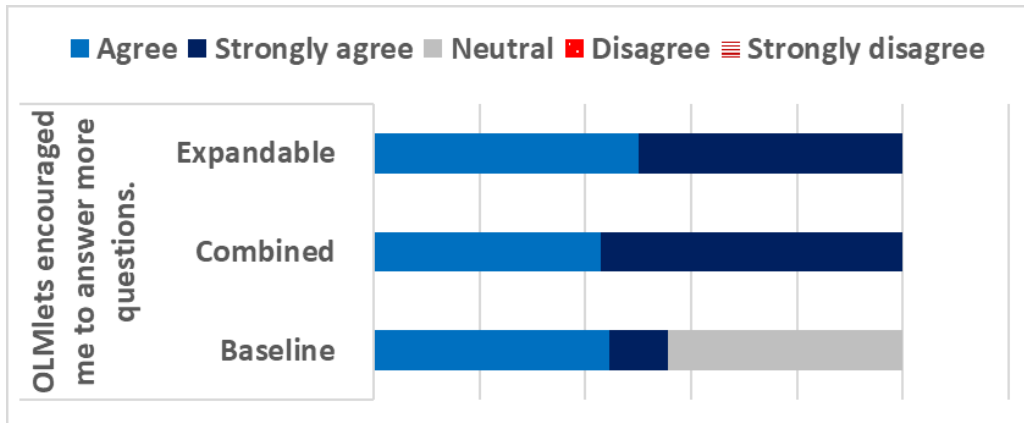


Figure 5-8. Post-questionnaire item of learning section where different views

### Students' perception on confidence after OLMlets use

Table 5-5. The mean rank for each group on the post-questionnaire items related to their confidence.

	Baseline	Combined	Expandable	$X^2(2)$	$P$
I am good at assessing my confidence level.	10.39	8.36	16.83	6.80	0.033
I try to increase my knowledge when the system shows that I had low level of knowledge and I feel that my confidence is low.	9.33	14.00	11.83	2.33	0.312
I try to increase my knowledge when the system shows that I had low level of knowledge and I feel that my confidence is high.	9.94	14.64	10.17	2.82	0.244

I felt my confidence increase when the system shows that I had a high level of knowledge.	11.06	10.36	13.50	1.05	0.592
I felt my confidence decrease when the system shows that I had low level of knowledge.	9.00	14.64	11.58	3.26	0.196
<b>Feeling that my confidence level matched my knowledge level encouraged me solve more questions.</b>	<b>7.44</b>	<b>11.86</b>	<b>17.17</b>	<b>9.06</b>	<b>0.011</b>

Two items on the post-questionnaire related to participants' confidence revealed significant differences '*I am good at assessing my confidence level*', using the Kruskal-Wallis test ( $\chi^2(2) = 6.80, p = .033$ ). The mean rank for both treatment conditions (expandable model = 16.83, combined model = 10.39) were higher than for the baseline condition (mean rank = 8.36). Participants from the expandable model who could view the alignment using the two options combined and expanded believed that they had good knowledge monitoring skills, compared to both the baseline and the combined models. The second significant item was '*Feeling that my confidence level matches my knowledge level encourages me to answer more questions*', ( $\chi^2(2) = 9.06, p = .011$ ). In this item, both treatment conditions felt that aligning their confidence to their performance encouraged them to answer more questions compared to the baseline condition, who had different views. Figure 5-9 shows the percentage distribution of participants' perception about the two items of the questionnaire that revealed significant difference related to students confidence after OLMlets use.

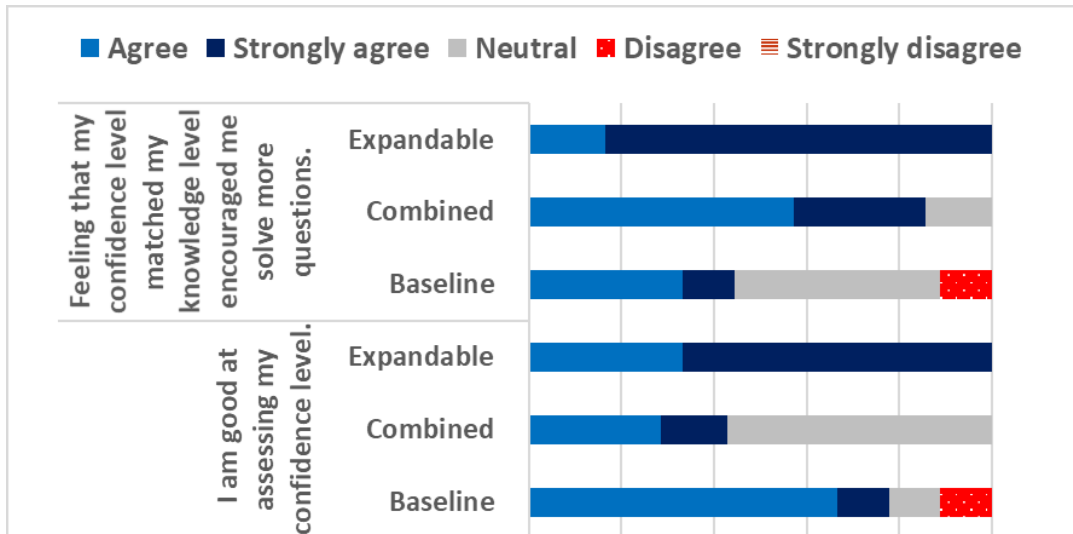


Figure 5-9. Post-questionnaire items for confidence section where different views were found after OLMlets use.

### **Students' perception on visualising alignment in OLMlets**

There were no significant differences in questionnaire items (Table 5-6) related to how participants interpreted the alignment for the third section related to students' perception in visualising alignment in OLM for students in both treatment conditions (combined and expandable). This indicates that participants in both treatment conditions had similar interpretations of the information depicted by those visualisations.

*Table 5-6. Questionnaire items for both treatment groups about the interpretation of visualising alignment.*

	<b>Combined</b>	<b>Expanded</b>	<b><math>\chi^2(2)</math></b>	<b><i>P</i></b>
The system helped me to identify topics where I am under-confident.	6.79	7.25	0.06	0.805
The system helped me to identify topics where I am over-confident.	5.36	8.92	3.60	0.058
I can easily notice from the chart, when my confidence and my level of knowledge match completely.	8.21	5.58	2.06	0.151
I can easily notice from the chart, when my confidence and my level of knowledge mostly match.	7.21	6.75	0.06	0.805
I can easily notice from the chart, when my confidence and my level of knowledge match somewhat.	5.07	9.25	4.91	0.027
I can easily notice from the chart, when my confidence and my level of knowledge are completely different.	5.93	8.25	1.35	0.245
Being shown that there is a difference between my knowledge and confidence is helpful.	6.57	7.50	0.22	0.642
Seeing my model let me be more accurate in assessing my confidence level.	7.29	6.67	0.11	0.735
Being shown that there is a difference between my knowledge and confidence led me to solve more questions.	7.00	7.00	0.00	1.000

### ***Students' perception on expanding the model in OLMlets***

Participants from the expandable model condition agreed that expanding the model into two separated models was useful and helped them in their knowledge monitoring skills. Figure 5-10 illustrates the percentage of participants views from

the expandable model condition about their ability to expand the models (see Appendix 11).

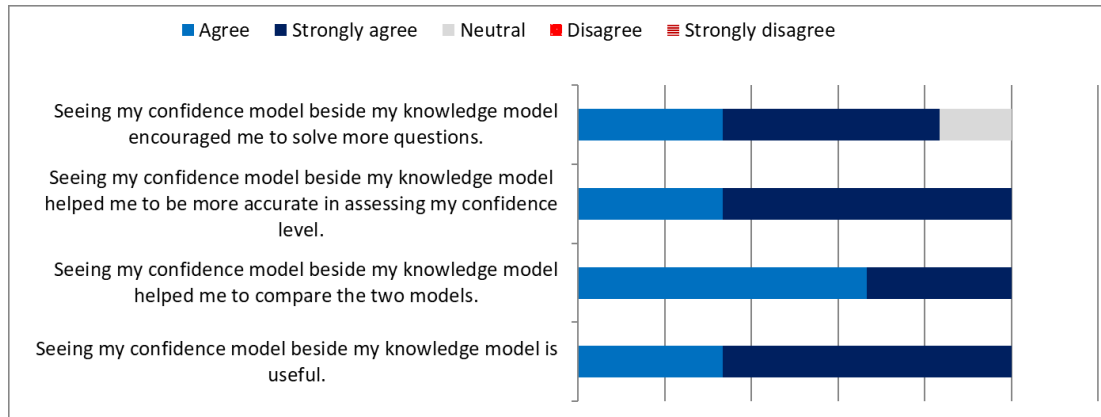


Figure 5-10. Students' views of the option to expand the models.

#### 5.4.3 Pre- and post-test of student knowledge

Table 5-7 presents the descriptive statistics for the pre- and post-tests for participants students who did not used the OLMlets and those did, together with the user sub-groups of OLMlets. The maximum score for both pre- and post-tests is 10.

Table 5-7. Descriptive statistics for pre- and post-tests for both non-OLM-users and OLM-users including the sub-groups of OLM-users.

Groups	n	Pre-test		Post-test	
		M (SD)	95% CI	M(SD)	95% CI
no-OLMlets	29	6.21 (2.53)	[5.25, 7.17]	5.52 (1.99)	[4.76, 6.28]
OLMlets	25	7.32 (2.51)	[6.28, 8.36]	6.52 (1.58)	[5.87, 7.17]
Baseline	9	8.25 (1.75)	[6.78, 9.72]	6.63 (1.06)	[5.74, 7.51]
Combined model	9	7.5 (2.20)	[5.66, 9.34]	6.5 (1.31)	[5.41, 7.59]
Expandable model	7	6.12 (3.87)	[2.11,10.23]	6.5 (2.07)	[4.32, 8.68]

To explore the effect of the OLMlets system on students' knowledge, a Mann-Whitney U-test was used to identify any significant differences on pre-test for



OLMlets users and non-OLMlets users. No significant difference was observed between the two groups using the Mann-Whitney U-test ( $U = 259$ ,  $p = .069$ ,  $r = .25$ ) indicating that both groups had comparable abilities in their knowledge at the start of the study. Although the post-test was normally distributed, a non-parametric test was used to allow for easily comparison between the findings from the pre- and the post-test. A non-parametric test (which is less powerful than a parametric test and does not have the stringent assumptions about the nature of data being analysed) showed that there was no significant difference in the post-test between the two groups (OLMlets users and non-OLMlets users) using the Mann-Whitney U-test ( $U = 251.5$ ,  $p = .051$ ,  $r = .26$ ). However, the test would tend to reveal differences when the nature of the data of the post-test was considered.

Pre- and post-tests for OLMlets users were normally distributed. Therefore, a one-way independent ANOVA was used to explore the impact of the visualisation condition between the three conditions of OLMlets users. There was no significant difference found in the pre-test using a one-way independent ANOVA ( $F(2) = 1.515$ ,  $p = .242$ ,  $\eta^2 = .12$ ), equal variance assumed. Further, no significant difference was observed between three visualisation conditions of OLMlets-users on the post-test, using a one-way independent ANOVA ( $F(2) = .265$ ,  $p = .770$ ,  $\eta^2 = .02$ ), equal variance assumed.

For all groups, student experience decreased in their performance in the post-test, which may be due to the level of difficulty of the questions compared to the pre-test. However, in this study, the focus is to compare the three conditions in each test, rather than comparing learning gained from pre- to post-test.

#### 5.4.4 Pre- and post-test of student confidence

A strong positive correlation was observed between students' confidence on the pre-test and students' pre-test score for all participants ( $r = .80, p < 0.001$ ). There was also a positive correlation between students' confidence on the post-test and students' performance on the post-test for all participants ( $r = .72, p < 0.001$ ). Although there was a decrease in correlation, it is still considered respectable. Table 5-8 presents the descriptive statistics for students' confidence on the pre- and post-test for participants who did not use the OLMlets and those who did, including the sub-groups of OLMlets-users.

*Table 5-8. Descriptive statistics for pre- and post-confidence for non-OLMlets users and OLMlets-users, including the three conditions of OLM-users.*

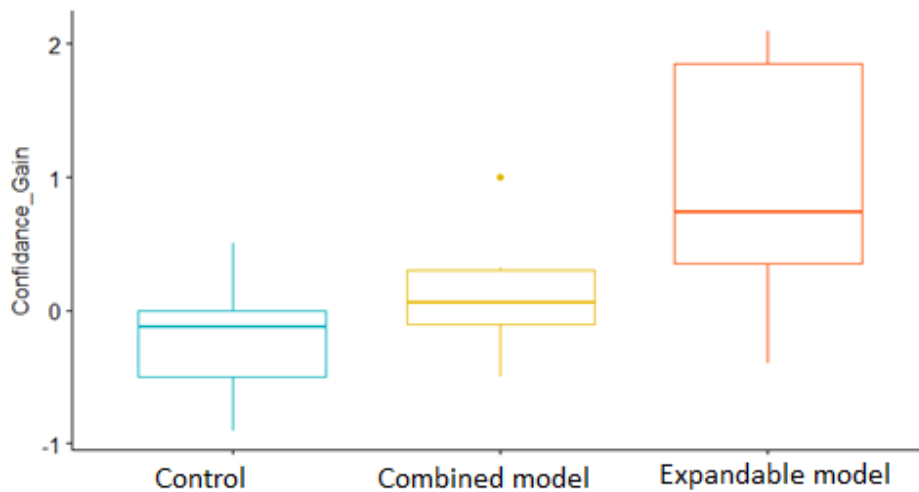
Groups	n	Pre-confidence		Post-confidence	
		M (SD)	95% CI	M (SD)	95% CI
<b>Non-OLMlets</b>	29	2.76 (0.64)	[2.52, 3.00]	2.82 (0.68)	[2.56, 3.08]
<b>OLMlets</b>	25	3.15 (0.65)	[2.89, 3.42]	3.40 (0.42)	[3.23, 3.58]
Control	9	3.41 (0.37)	[3.12, 3.69]	3.20 (0.45)	[2.85, 3.55]
Combined	9	3.24 (0.47)	[2.88, 3.60]	3.38 (0.34)	[3.12, 3.65]
Expandable	7	2.71 (0.94)	[1.85, 3.58]	3.69 (0.33)	[3.38, 4.00]

Although all students in the study volunteered to use the system and were distributed among the groups randomly, not all students actually logged to the system. This enabled an exploration of whether confidence differed between those

who were motivated to use the system and those who were not motivated to use the system. This was supported by the significant difference found in students' confidence in the pre-test using the Mann-Whitney U-test ( $U = 230$ ,  $p = .021$ ,  $r = .31$ ), with a mean rank of 22.93 for non-OLMlets users and 32.80 for OLMlets users, showing that those who used the OLMlets had higher confidence than those who did not use the OLMlets system. There was also a significant difference in confidence in the post-test, between participants who used the OLMlets and those who did not, indicated by the Mann-Whitney U-test ( $U = 176$ ,  $p = .001$ ,  $r = .44$ ) with a mean rank of 21.07 for those who did not use the OLMlets and 34.96 for those who did use the OLMlets. The post-test effect was stronger than that of the pre-test, which could reflect that using the OLMlets system might influence students' confidence. Those who used the OLMlets system experienced a small confidence gain ( $M = 0.25$ ,  $SD = 0.78$ , 95% CI = [-0.07, 0.57]) in contrast to those who did not use the OLMlets system, who experienced almost no confidence gain ( $M = 0.06$ ,  $SD = 0.76$ , 95% CI = [-0.23, 0.35]).

Analysing pre-confidence for the sub-groups of OLMlets users, no significant differences were found using the Kruskal-Wallis test ( $\chi^2(2) = 2.94$ ,  $p = .230$ ,  $\eta^2 = .12$ ) with a mean rank of 15.78 for the control group, 13.00 for the combined model group, and 9.43 for the expandable model group. Furthermore, no significant differences were found for students' post confidence for the sub-groups of OLMlets users using the Kruskal-Wallis test ( $\chi^2(2) = 5.28$ ,  $p = .068$ ,  $\eta^2 = .22$ ), with a mean rank of 9.44 for the control group, 12.67 for the combined model group, and 18 for the expandable model group.

The confidence gain for each of the sub-groups of OLMlets users was normally distributed, based on the results of the Shapiro-Wilk test ( $p > .05$ ); thus, a one-way independent ANOVA was used to compare students' confidence gain between the three visualisation conditions. However, the Levene's F test revealed that the assumption of homogeneity of variance was violated ( $p = .007$ ). As a result, the Welch's F test was used ( $F(2) = 4.52$ ,  $p = .034$ ,  $\eta^2 = .39$ ). Figure 5-10 shows the box-plot for students' confidence gain between the three visualisation conditions of OLMlets users, where both treatment conditions showed an increase in their confidence gain. The combined model condition had ( $M = 0.14$ ,  $SD = 0.42$ ,  $95\% \text{ CI} = [-0.18, 0.46]$ ) and the expandable model condition ( $M = 0.98$ ,  $SD = 0.98$ ,  $95\% \text{ CI} = [0.07, 1.9]$ ). In contrast, the control condition had a decrease in their confidence ( $M = -0.21$ ,  $SD = 0.46$ ,  $95\% \text{ CI} = [-0.56, 0.14]$ ), shown in Figure 5-11 which illustrates that the control condition (where the alignment information was hidden) was not effective, as both treatment conditions were successful in enhancing students' confidence.



*Figure 5-11. Box plots with error bars for students' confidence gain for the three groups of OLMlets users.*

In Figure 5-11, a dot above the error bar of the combined model is shown to indicate an outlier. The analysis was repeated with and without the outlier, and the results remained the same; thus, the outlier was included in the analysis because it reflects that one participant from the combined model had very high increase in their confidence gain compared to others in the combined model.

The post-hoc analysis used the Bonferroni method to identify which groups were significantly different (Table 5-9). As shown in the table, a significant difference was observed between the control condition and the expandable model condition with ( $Z = -2.75, p = 0.006$ ).

Table 5-9. Post-hoc analysis using Bonferroni method for students' confidence gain with significant p-value indicated in bold text.

	Z	P-value
Control - Combined	-1.41	0.158
Control - Expandable	<b>-2.75</b>	<b>0.006</b>
Combined - Expandable	-1.43	0.153

Exploring relationships between students' confidence and students' performance on tests for each of the sub-groups of OLMlets users, there was a positive relationship between students' pre-test confidence and pre-test performance ( $r = .73$ ,  $p < 0.001$ ). There was also positive relationship between students' post-test confidence and post-test performance for OLMlets users ( $r = .61$ ,  $p = .001$ ). The reduced correlation could be due to the level of difficulty of the post-test. However, there was no relationship between confidence pre-test and post-test for OLM-users ( $r = -.01$ ,  $p = .948$ ). There was also no relationship between pre-test or post-test performance for OLMlets users ( $r = .01$ ,  $p = .967$ ). This lack of relationship could be owing to the level of difficulty of the post-test, where participants had lower scores in the post-test than the pre-test (see Table 5-7), which could influence their confidence.

#### 5.4.5 System use

Based on the system logs data, Table 5-10 shows the descriptive statistics for students' system use for the three visualisation conditions of OLMlets-users. The number of times the participants viewed the model in the expandable model

group was counted when the student clicked to view the model, regardless of number of times they expanded the model.

*Table 5-10. Descriptive statistics for students' system use for OLM-users.*

<b>Groups</b>	<b>No. Q. Answered</b>		<b>No. Model Views</b>	
	<b><i>n</i></b>	<b><i>M</i> (SD) 95% CI</b>	<b><i>M</i> (SD)</b>	<b>95% CI</b>
<b>Control</b>	9	23.00 (6.20) [18.23, 27.70]	3.11 (5.25)	[-0.93, 7.15]
<b>Combined model</b>	9	37.56 (25.26) [18.14, 56.98]	3.78 (5.09)	[-0.14, 7.69]
<b>Expandable Model</b>	7	59.29 (23.75) [37.32, 81.25]	8.86 (10.33)	[-0.70, 18.42]

Figure 5-12 shows the box-plot for the three conditions for the number of questions answered. Using the Kruskal-Wallis test, there was a significant difference observed between conditions ( $\chi^2(2) = 7.62, p = .022, \eta^2 = .32$ ). There were some outliers that I have included in the analysis indicated in Figure 5-12 (as dots below and above the error bars). However, using a non-parametric test is more robust for such outliers. Including the outliers in the analysis provides more information about the behaviour of some participants who were able to complete more or less activities than others from the same group.

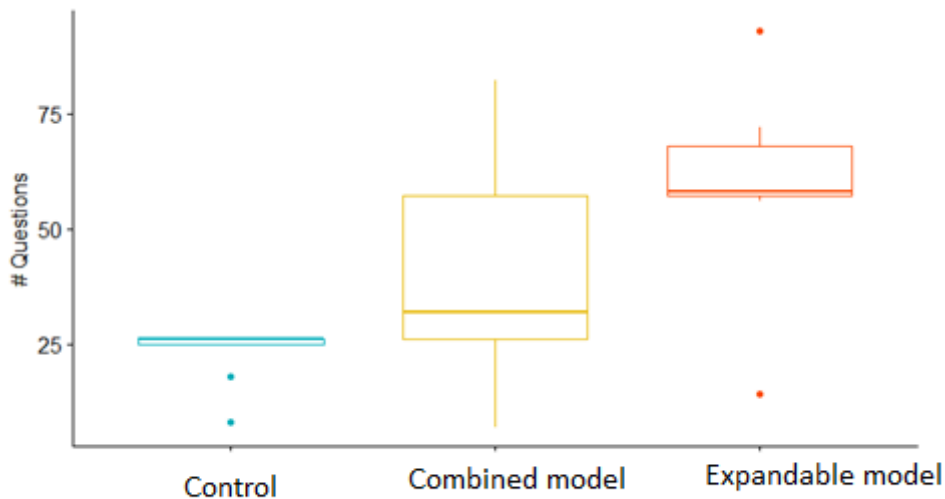


Figure 5-12. Box plots with error bars for the three groups of OLMlets users for number of questions answered.

The post-hoc analysis shown in Table 5-11 suggests that the difference occurred between the control condition and the expandable model condition. This indicates that visualising alignment motivated students to interact with the system and carry out more learning activity, compared with the control condition who had fewer questions solved compared to the other conditions.

Table 5-11. Post-hoc analysis using Bonferroni method for number of questions answered with significant p-value indicated in bold text.

	Z	P-value
Control - Combined	-1.42	0.16
Control - Expandable	<b>-2.76</b>	<b>0.006</b>
Combined - Expandable	-1.43	0.15



Students' behaviour in viewing the model also revealed significant differences between the three conditions using the Kruskal-Wallis test ( $\chi^2(2) = 8.74, p = .013, \eta^2 = .36$ ). Figure 5-13 shows the box-plot for the number of times the model was viewed for the three conditions.

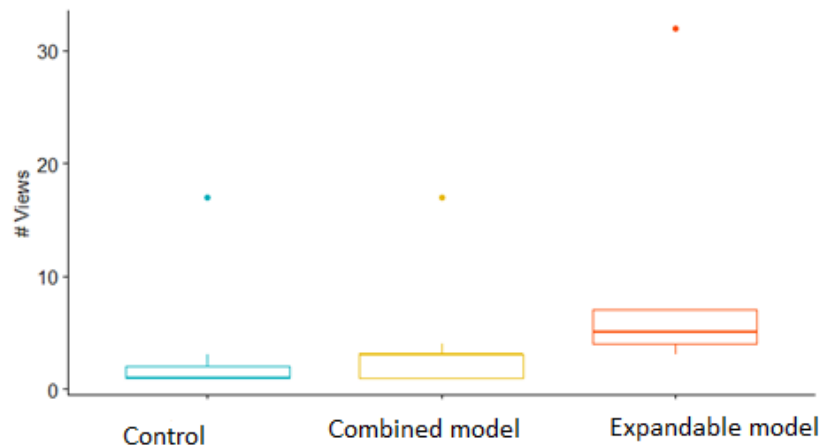


Figure 5-13. Box plots with error bars for the three groups of OLMlets users for number of times the model was viewed.

The post-hoc analysis using Bonferroni method for multiple comparisons is shown in Table 5-12.

Table 5-12. Post-hoc analysis using Bonferroni method for number of times the model was viewed with significant p-value indicated in bold text.

	Z	P-value
Control - Combined	-0.83	0.404
Control - Expandable	<b>-2.90</b>	<b>0.011</b>
Combined - Expandable	<b>-2.12</b>	<b>0.034</b>

The post-hoc analysis suggested that the significant difference occurred between the control condition and the expandable model condition. A significant difference also occurred between the combined model condition and the expandable model condition. This indicates that participants in the expandable model group were more encouraged to view the model and to know how their confidence was aligned to their knowledge, based on the system evaluation. Based on the system log data, participants expanded the model an average of 4.86 times ( $SD = 5.79$ ) while viewing their OLM.

#### **5.4.6 Relationship of system use to confidence**

To explore how OLMlets use affected confidence, a Pearson's correlation of student confidence in the pre-test with number of questions answered was explored suggesting, no-relationship ( $r = -.16, p = .437$ ). Further, there was no relationship between number of questions answered and students' performance in the pre-test ( $r = -.21, p = .303$ ). However, a moderate positive relationship was observed between students' confidence in the post-test and number of questions answered ( $r = .51, p = .009$ ). This suggest that interaction within OLMlets positively affected student confidence.

### **5.5 Discussion**

Although, there was no significant difference between the sub-groups of OLMlets users in students' performance, there was a large effect for those participants who used the treatment condition ( $\eta^2 = .39$ ) in increasing student confidence, particularly for participants from the expandable model condition. Further,

confidence was related to performance in both the pre- and post-tests, confirming prior results (Jackson and Kleitman, 2014) within a new instructional domain.

The results indicated that seeing the misalignment between their confidence and performance had an influence on students' confidence, which in turn influenced their behaviour in using the system. Based on the theory in this thesis that students who can view the alignment information between their confidence and their performance, may work hard when misalignment occurred (Kulhavy and Stock 1989). Based on the logged data of interaction with the system, the findings show that both treatment conditions answered more questions than the control condition, providing evidence that supports the theory of this thesis. Moreover, participants from the expandable model condition who had the ability to expand the model to view two separated models were more motivated to view the OLM than the other conditions. This indicates that students benefited from the two approaches to viewing the model (combined and expanded).

Based on the findings, it is suggested that visualising alignment had an impact on students' interaction with the system, which in turn affected their confidence; thus, confirming the hypothesis of this study. Prior research showed that allowing students to compare directly between students' confidence in the correctness of the answers and the computer evaluation about students performance could resolve the misalignment between students' confidence and students' performance through negotiation (Bull and Pain, 1995; Kerly and Bull, 2008). This study shows that simple OLMs can stimulate students' confidence and influence their learning

behaviour in using the system, such as answering more questions of the systems' automated assessment and increasing their viewing of the model.

# **Chapter 6 : EVALUATING THE EFFECT OF VISUALISING ALIGNMENT ON STUDENTS' BIAS OF CONFIDENCE JUDGEMENT**

---

## **6.1 Overview**

In this thesis, previous studies have investigated the impact of visualising alignment of students' confidence using paper-based pre- and post-tests. Accordingly, findings suggest that visualising alignment improved students' confidence (this was compared to students from the control condition who experienced a reduction in their confidence). Further, findings show that visualising alignment had an impact on students' behaviour in using the system, whereby students from the expandable model had a greater number of questions answered and viewed the model more. In this chapter, I expand my research to focus on the effect of visualising alignment to students' bias of confidence judgement while answering the system-automated assessment. In this case, confidence judgement is the act of students judging their confidence for each answer they provide. Students' bias of confidence judgement is the degree to which students are under-confident or over-confident, where a bias of 0 indicates that students aligned their confidence

in the correctness of the answers to the computer evaluation of students' performance. This study hypothesises that visualising alignment will reduce students' bias of confidence judgement resulting in more accurate confidence judgement when students can see both their beliefs about their confidence that the answers are correct and their performance in the answers given.

This and the following chapter present two studies that were conducted with the same cohort of undergraduate students over two semesters, during which they followed different computer science courses each semester. In this study, there were two visualisation conditions based on whether or not the alignment information was visualised in the OLMlets system. In the following study, the learner model representation is based on how the alignment information is presented (textual vs. graphical), which addresses the third research question of this thesis. Semi-structured interviews were conducted at the end of the second study to explore students' perspectives of their experience for each presentation of the learner model across both studies, and how these presentations supported students' bias of confidence judgement. The study in this chapter answers the research question of this thesis, which is 'how does visualising alignment support students' confidence judgement?'

## **6.2 Introduction**

Research indicates that students who perform poorly in their learning experience deficiencies in their knowledge monitoring skills, compared to high-achieving students (Hacker et al., 2000). Further, research has shown that there is a relationship between students' knowledge monitoring and their performance (Tobias and

Everson, 2002). Experimental studies suggest that where students are over-confident in estimating their knowledge, this could lead to them suspending their learning process to ensure that they have mastered the topic (Koriat, 1997; Kellemen, 2000). It is also noted that motivation could be an important factor to enable students to engage in metacognitive thinking to regulate their own knowledge (Zimmerman, 1990). Therefore, Intelligent Tutoring Systems (ITSs) could play a vital role in motivating students to interact with the system, as well as being used to enable monitoring processes. This is achieved by visualising the underlying user model to create OLMs.

Students providing answers to questions may perform correctly, but this correct answer may represent either solid knowledge or a random guess. Students with good knowledge monitoring skills can accurately align their own confidence in the correctness of their answers to their actual performance. This chapter extends the study described in the previous chapter, to evaluate the theory of how visualising alignment in OLMs can influence students' confidence judgements and their behaviour while using the system. In line with other researchers who measure students' confidence judgement, one popular approach is to ask students '*how confident are you that your answer is correct?*' (Jackson and Kleitman, 2014). In this study, the self-report answer follows after each question completed by students in the system's automated assessment for both groups. Students from the control condition group could view their performance in the answers given only in their learner model. However, students from the treatment condition group could inspect the alignment between their performance and their confidence in the

correctness of the answers given. Kulhavy and Stock (1989) argue that when misalignment is detected between students' confidence and their performance, the student may seek to resolve this misalignment. Thus, visualising alignment between the computer's evaluation and students' confidence in their performance may have implications for students' metacognitive skills and how students use OLMs.

One of the most used measures to explore students' tendency toward either over- or under-confidence regarding their knowledge is the bias of confidence judgement (Schraw 2009), discussed in the first chapter of this thesis. The bias measure reflects how students could align their performance with their own confidence that their answers are correct (see Equation 3) (Schraw 2009). The greater the misalignment, the larger the score on the bias index. The bias score reflects that the student is over-confident when the bias score is positive and under-confident when the bias score is negative. A bias score of 0 indicates that the students aligned their performance with their confidence.

$$\text{Bias} = \frac{\sum_1^n (\text{confidence} - \text{performance})}{n} \quad [3]$$

Thus, this type of bias represents the accuracy—or alignment—of a student's self-assessed confidence, which is referred to as confidence judgement in the system evaluation.

Accordingly, the hypotheses of this study are as follows:

**H<sub>1</sub>:** Visualising alignment will reduce students' bias in their confidence judgement.



**H<sub>2</sub>:** Visualising alignment will motivate students to answer more questions and to view the model more.

## **6.3 Method**

The OLMlets system was evaluated using a real classroom environment. An overview of the system design and architecture of the OLMlets system is described earlier in this thesis in Chapter 3. This section details the study's participants, design of the questions integrated into OLMlets to fit the course curriculum, data collection, measures, and analysis.

### **6.3.1 Participants**

Participants were second-year undergraduate students in the School of Electronics, Electrical and Systems Engineering at the University of Birmingham in the United Kingdom. There were 46 students in the class. Of those 46 students, 37 logged in to the system. Students were assigned an identifier number from S1 to S46 to provide their identity anonymously while showing their individual use of the system. Using the system was voluntary, and students used it in their own time.

As with all study procedures in this thesis, students who did not view the model or viewed it only once were excluded from the analysis ( $n = 4$ ). In addition, students who attempted less than 10 questions were excluded ( $n = 1$ ). Thus, five students were excluded from those ( $n = 37$ ). Students from each condition were classified for analysis into groups based on the median score of their performance at the time of answering the first set of questions in OLMlets ( $Mdn = 62\%$ ).

Students below the median were classified as low-achieving students and those at or above the median were classified as high-achieving students. Table 6-1 shows the number of students who participated in the study (N = 32) assigned to two conditions and classified by their ability level.

*Table 6-1. Number of participants in each condition classified by students' ability level.*

Ability level	Median	Conditions	
		Control	Expandable
Low-achieving	< 62%	7	11
High-achieving	>= 62%	7	7
<b>Total</b>		14	18

### 6.3.2 Study design

The study investigated the impact of visualising alignment on students' bias of confidence judgement and the students' behaviour in viewing the model and the number of questions answered.

To explore perceptions of the learner model communication approach, a mixed-methods explanatory design was used (Creswell and Plano Clark, 2007), where quantitative data regarding student behaviours, self-perception, and knowledge were complemented with more open-ended data about student experiences.

The study was conducted in the first semester of the second year of undergraduate study, and took place over a two-week period near the end of term when students volunteered to use the system to prepare for their course tests.

There were two sets of revision questions. Set 2 was uploaded to OLMlets one week after uploading Set 1. This was done to explore the practise effect on students' confidence judgments from Set 1 to Set 2.

The study had two visualisation conditions, based on whether or not the alignment information was visualised in the OLMlets system. Participants were randomly assigned to a condition. Participants in each condition were later classified into two levels (low- or high-achieving students). The first condition was the control condition, where participants used the standard skill meter; thus, the alignment information was hidden in the OLMlets system, see Figure 6-1.

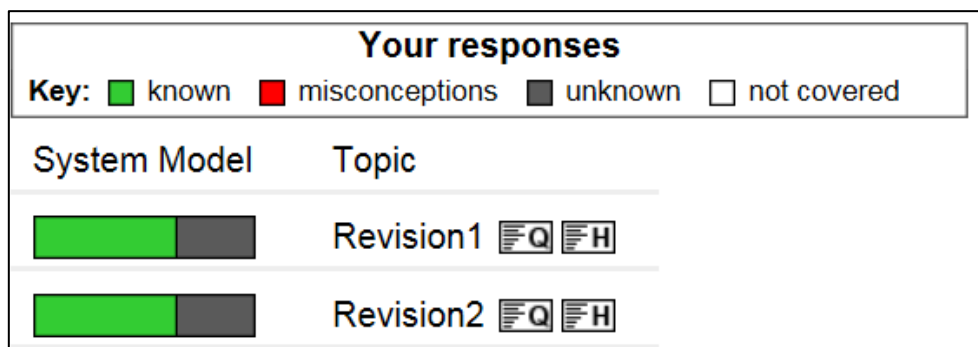


Figure 6-1. The control condition within OLMlets specified by a participant from this study.

The treatment condition was the expandable model condition, where participants could see the alignment information from their learner model using two methods: combined (Figure 6-2) or expanded (Figure 6-3). Both Figures 6-1 and 6-2 are learner models of students who participated in this study. Participants assessed their confidence after answering each question, regardless of condition.

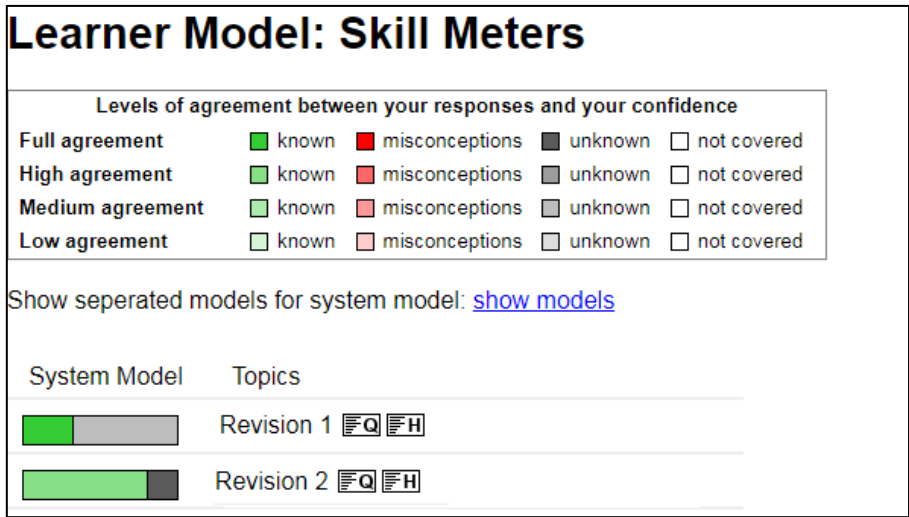


Figure 6-2. The expandable model condition in the combined view.

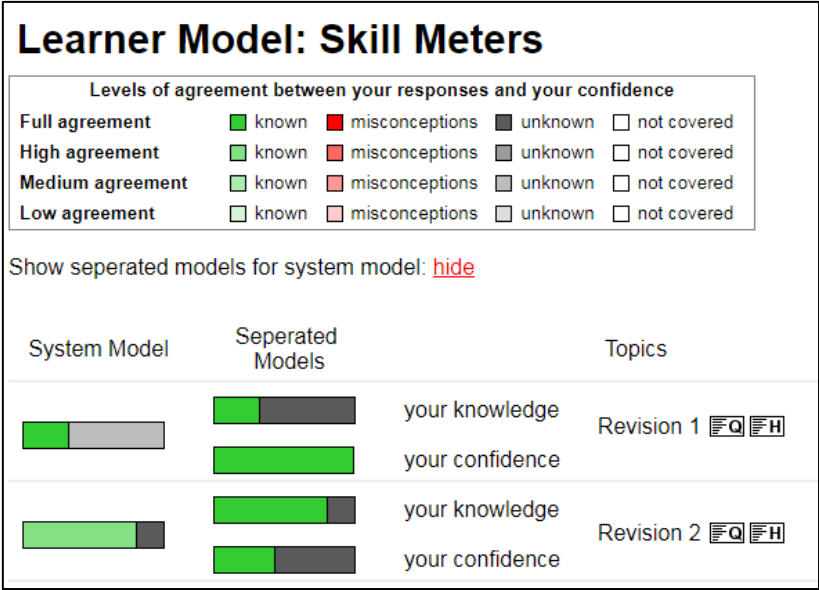


Figure 6-3. The expandable model in both views (a) combined and (b) expanded specified by participant from this study.

To gain a broad sense of how participants responded to different conditions, participants were recruited from the same cohort so that it would be possible to compare their perceptions of the open learner model communication approaches

across a variety of conditions. To achieve this goal, a mixed-method explanatory design was used (Creswell and Plano Clark, 2007), where quantitative data regarding student behaviours, self-perception, and knowledge were complemented with more open-ended data about student experiences.

### **6.3.3 System training**

Each participant received a 15-minute presentation, including a live demonstration, to explain the system and its features. During this presentation, participants were shown each of the OLMlets study conditions. On reflection, revealing the alternative conditions might have been sufficient for participants to have become sensitized to the aim of the experiment. In other words, knowing the alternative conditions might have heightened their response (either positive or negative) to the condition that they experienced. If the study was to be repeated, it might be sensible for the presentation to only include the condition which participants would experience. In defence of the approach taken, it was felt that it was important for participants to appreciate how the system operated and this involved them seeing the ways in which it could be used. A user manual was distributed to all participants, dependent upon which condition the student was assigned. The manual contained a visual step-by-step guide to system features that included screenshots to help understand the system. Each participant had a username and password.

### **6.3.4 Design of the questions integrated into OLMlets (Java Programming)**

The study was based on a programming course (*Introduction to Java Programming*). The OLMlets system was used as a preparation tool to enable participants

to prepare for their final exam. Thus, two sets of review questions that covered the course curriculum were integrated in OLMlets as two sets of revisions (Revision 1 and Revision 2), as shown in Figures 6-1, 6-2 and 6-3. In each revision set, there were 25 multiple-choice questions, and for each, the participants were required to self-assess their confidence that the answer they provided was correct.

Prior research has shown that question difficulty can affect student confidence judgements, with students tending to be over-confident when questions are more difficult (Nietfeld et al., 2005; Schraw and Dennison, 1994). To help control the effect of the level of the difficulty of the questions integrated to OLMlets, Bloom's Taxonomy was used as a proxy for question difficulty, because questions requiring knowledge of specific principles or facts tend to be answered correctly more than questions requiring both knowledge of the principles and how to apply them in a new situation (Bloom et al., 1956).

Thus, questions in OLMlets were categorised according to Bloom's Taxonomy, from the simple recall of information and understanding of concepts to more complex questions that require evaluation skills. I have performed the categorisation of the questions based on Bloom by following the criteria for each category. The taxonomy contains six categories (Bloom et al., 1956), as follows:

- Knowledge: The recall of major principles and theories.
- Comprehension: The ability to understand basic statements.

- Application: The ability to remember and apply principles, theories, and ideas.
- Analysis: The ability to breakdown the statement for clarification.
- Synthesis: The ability to put together elements to form a whole.
- Evaluation: The ability to assess and judge for a given purpose.

To improve upon this categorisation, the course instructor rated the difficulty of each question on a scale, from 1 (very easy) to 5 (most difficult). This rating was performed independently of the Bloom's Taxonomy categorisation. These two methods of categorising the questions meant that all questions had been assigned to a category from Bloom's Taxonomy and given an instructor-determined difficulty level.

Table 6-2 shows the number of questions in each Bloom category and the difficulty rating, based on the instructors' perception. The difficulty rating of each Bloom category is the average score of the level of difficulty of questions (rated by the instructor) that belong to the same Bloom category. For example, there were 14 questions in Set 1 that belonged to the Bloom category 'knowledge', and the average difficulty level of the 14 questions was 3 (see Table 6-2). This ensured that both Set 1 and Set 2 ranges were within the same difficulty level and classified within the same Bloom categories. Thus, participants' confidence judgement would not be influenced by the level of difficulty of the questions from Set 1 to Set 2.

Table 6-2. Number of questions, difficulty rating based on instructor's perception for the course Java programming.

	Knowledge	Comprehension	Application	Analysis	Synthesis	Evaluation
<b>Set 1</b>						
<b>Number of questions (N = 25)</b>	14	4	3	4	0	0
<b>Average difficulty rate (1 → 5)</b>	3	3	3	4	-	-
<b>Set 2</b>						
<b>Number of questions (N= 25)</b>	15	4	2	4	0	0
<b>Average difficulty rate (1→ 5)</b>	2	3	3	3	-	-

The following examples show questions integrated into OLMlets for each category of Bloom's Taxonomy, except for the last two categories (synthesis and evaluation). These two levels were considered beyond the scope of the multiple-choice assessment.

**Knowledge:**

Which of these keywords is used to refer to a member of base class from a sub class?

- a) upper
- b) super
- c) this
- d) extends

**Comprehension:**

What pattern does the Flow layout manager use to add components to a container?

- a) Left to right, top to bottom.



- b) Bottom to top, right to left.
- c) Top to bottom, centred in each row.
- d) Any order

**Application:**

The method \_\_\_\_ sets the font (Helvetica, 20-point bold) in component C

- a) c.setFont(new Font("Helvetica", Font.bold, 20))
- b) c.setFont(new Font ("helvetica", BOLD, 20))
- c) c.setFont(Font("Helvetica", Font.BOLD, 20))
- d) c.setFont(new Font("Helvetica", Font.BOLD, 20))

**Analysis:**

Show the output of the following code:

```
Import javax.swing.*;
Public class Test {
Public static void main (String[ ] args) {
JButton jbtoK = new JButton("OK");
System.out.print (jbtoK.isVisible()+ ",");
JFrame frame = new JFrame();
System.out.println(frame.isVisible());
}}
```

- a) true, true
- b) true, false
- c) false, true
- d) false, false

**6.3.5 Data collection**

The study yielded both quantitative and qualitative data. The quantitative data was derived from the system's logging of participants' interactions with OLMlets, and included performance on learning tasks, confidence ratings for their

assessment, the number of questions answered, and the number of times the model was viewed. The qualitative data was derived from semi-structured interviews conducted after the end of both studies presented in this and the following chapters. Quotations from participants were used to explain relationships found in the quantitative data.

### **6.3.6 Measurement and quantitative analysis procedures**

Participant confidence and performance were derived from logs of the first attempt at a question (N = 25) for each question set. Performance was calculated based on the answer being correct or incorrect.

Bias of participants' confidence judgement was used as a proxy for knowledge monitoring skills, where it showed the direction of the individual's ability to accurately estimate their performance. Bias was calculated as a unitised difference of the confidence judgement and the performance across all questions answered in each of Set 1 and Set 2, using Equation [4], (Schraw 2009). The bias score is always between -1 and +1. The sign magnitude for the bias score is negative for under-confidence, positive for over-confidence, and a score of 0 indicates that performance was aligned with confidence judgement.

$$\text{Bias} = \frac{\sum_1^n(\text{confidence} - \text{performance})}{n} \quad [4]$$

Based on Equation 1, a bias score of 1 was obtained when the student failed to answer their knowledge assessment correctly but they were confident that the answer was correct. Conversely, a bias score of -1 is obtained when the student never fails in their knowledge assessment but they are unconfident that the

answer is correct for the whole assessment. The score of 0 (null score) is obtained when the student aligns their confidence to their performance.

Two independent variables were used. The first independent variable was visualisation condition (control, expandable model). The second independent variable was student ability level (low-achieving, high-achieving). Four dependent variables were measured. The first was the bias of students' confidence judgement for Set 1 and Set 2 (denoted as bias 1 and bias 2). The second was student performance on Set 1 and Set 2. This performance is based on the correctness of the answers given by students (performance 1 and performance 2). Two dependent variables were used to capture system use; these were the number of questions answered and the number of times the model was viewed.

A two-way ANOVA was used to determine if there was an interaction between the two independent variables (visualisation condition and student ability level) on the dependent variable (bias of confidence judgment) for both Set 1 and Set 2. An independent t-test was used to identify significant differences between conditions (i.e., control and expandable). Further, a paired sample t-test was used to identify differences in student performance across question sets. The threshold for statistical significance was 0.05. Mann-Whitney U-tests were used to identify differences between conditions when the data were not normally distributed. Bonferroni correction was employed to control for multiple comparisons, when needed.

In this study, the effect size was based on ( $\eta^2$ ) for ANOVA tests. Partial  $\eta^2$  was calculated and reported, based on Cohen's guidelines for effect size (Cohen, 1988). An effect size of 0.01 indicates a small effect, 0.06 a medium effect, and 0.138 a large effect. The effect size of Cohen's  $d$  was calculated for t-tests and followed the guidelines of 0.2 as small effect, 0.5 for medium effect, and 0.8 for large effect (Cohen, 1988). The effect size of non-parametric tests ( $r$ ) is indicated as 0.1 as small effect, 0.3 as medium effect, and 0.5 as large effect (Cohen, 1988).

### **6.3.7 Qualitative analysis procedures.**

Qualitative data were collected through semi-structured interviews after both studies had ended. Because the same students participated in both studies (Chapters 6 and 7), information about students' perceptions of their experiences in using the system across both semesters was collected to understand their use and reactions to the different methods for representing their learner model better. During these interviews, students were asked questions such as the following:

- *'Can you describe for me how you used the system?'*;
- *'You were asked to assess your confidence after each question, what did you think of this?'*;
- *'How did you use the information that was given to you about your confidence and performance in solving the questions?'* (see Appendix 13).

All interviews were recorded and transcribed. Data was then categorised into themes. There were five themes generated from the transcribed data: Exam

preparation, understanding a topic, revision strategy, confidence in understanding, and visualisation preference. These themes were used to understand student experiences more easily and explain the behaviours that were observed in the system log files. Although there was a long time between the study presented in this chapter and the interviews, students were still able to recall and explore their experiences between the visualisation presented in the first study and the visualisation they had in the second study. Consequently, suitable quotations were used to augment the quantitative results. Appendix 14 includes all students' transcribed data, where quotes were divided into themes. Those quotes used to augment the quantitative data in the thesis have been indicated in bold text.

## **6.4 Results**

### **6.4.1 Normality of data**

A Shapiro-Wilk test output for the bias 1 and bias 2 measures for both student ability levels in each condition showed that data were normally distributed ( $p > .05$ ). Student performance in Sets 1 and 2 were also normally distributed ( $p > .05$ ), except for that of Set 2 of low-achieving students from the control condition ( $p = .005$ ). The number of times students viewed the model in Sets 1 and 2 was also normally distributed ( $p > .05$ ). The number of questions answered was normally distributed for low-achieving students, but not normally distributed for high-achieving students for both groups. Appendix 12 shows students logs data for all students for the data used in this study.

#### 6.4.2 Students' bias of confidence judgement

Two-way ANOVA tests revealed no interaction between the two independent variables; visualisation condition and students' ability level on bias 1 [ $F(1) = .332, p = .569, \eta^2 = .012$ ] or bias 2 [ $F(1) = .015, p = .904, \eta^2 = .001$ ]. The interaction graph for bias 1 shows that the graphs intersect (Figure 6-4). However, the p-value of the interaction shows the interaction is not significant (visualisation condition \* ability level > .05). The apparent intersection is due to high variability in the data. The result shows that there was a main effect of the independent variable visualisation condition on the dependent variable with respect to students' bias of confidence judgement (bias 1) [ $F(1) = 6.625, p = .016, \eta^2 = .191$ ].

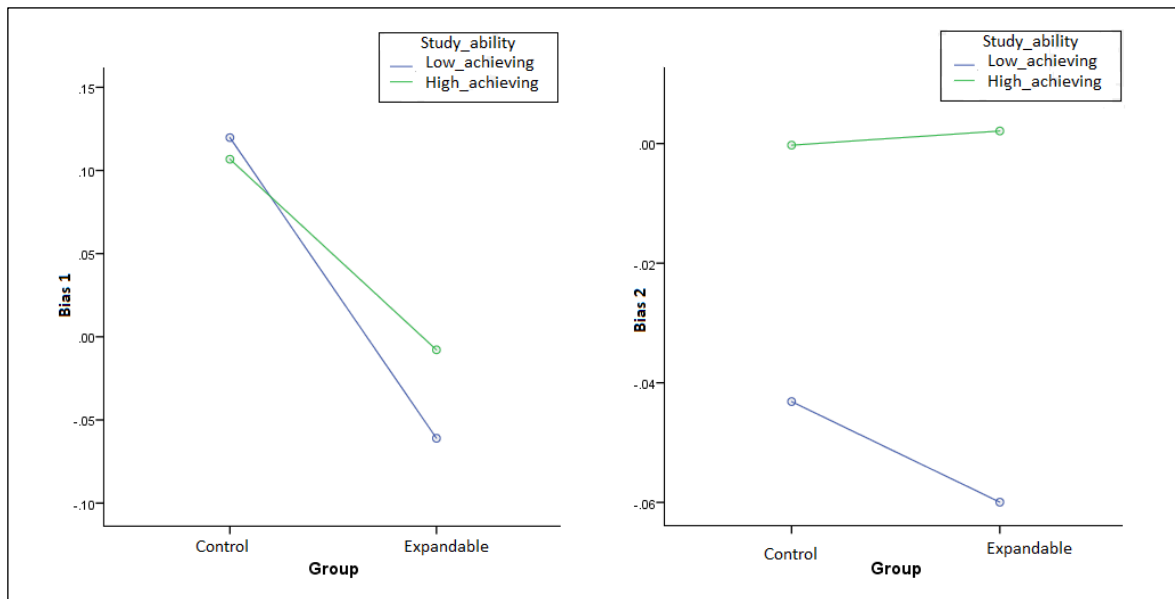


Figure 6-4. Two-way ANOVA interaction graphs between the two independent variables for bias 1 (a) and bias 2 (b).

An independent t-test (equal variance assumed), showed a significant difference between the control condition and the expandable model condition ( $t(30) = 2.77$ ,

$p = .009$ ,  $d = 1.02$ ) across all students. Students from the control group judged their confidence (bias 1;  $M = .11$ ,  $SD = .14$ ) less accurately than those from the expandable model condition ( $M = -.04$ ,  $SD = .16$ ). This shows that visualising alignment benefited students by helping them align their confidence with their performance. The significant difference was not observed in Set 2 using independent t-test (equal variance not assumed) ( $t(25.448) = .145$ ,  $p = .992$ ,  $d = .05$ ) across all students. Students from both groups had a similar ability to assess their confidence in their performance with ( $M = -.02$ ,  $SD = .21$ ) for the control group and ( $M = -.03$ ,  $SD = .20$ ) for the expandable model group. This shows that practising confidence judgement during question Set 1 may have impacted positively on the control group, who showed more accurate confidence judgement in Set 2 (as indicated by the bias scores), whereas the expandable model group had stable accurate confidence judgement in both Sets 1 and 2.

To address how different student background preparation might influence the bias in their confidence judgement, I separated the analysis based on student ability level for each condition.

***Low-achieving students' bias of confidence judgement by condition:***

A Bonferroni-corrected independent t-test (with equal variance assumed) showed a trend in participant bias for low-achieving students between the expandable model condition and the control condition for Set 1 ( $t(16) = 2.16$ ,  $p = .092$ ,  $d = 1.06$ ). Low-achieving students' bias of confidence judgement from the expandable model group seems more accurate for Set 1 ( $M = -0.06$ ,  $SD = 0.19$ ) than the

control group, who showed over-confidence ( $M = 0.12$ ,  $SD = 0.15$ ). This shows that visualising alignment might support students to align their confidence with their performance.

This potential difference in low-achieving student bias between the two conditions in Set 1 could be because students from the expandable model condition, who could inspect the alignment between their confidence and their performance, were taking their self-confidence assessment more seriously and trying to align their confidence to their performance.

This interpretation is supported by students' perspectives on their use of the system. A student from the control condition claimed that: *'I did [accurate self-assessment] the first few times and then afterwards I just wanted to do the questions so I didn't actually pay attention to that [confidence assessment]'* (S20). The log data for this student, S20, showed they made accurate confidence judgements (i.e., bias approaching the ideal of zero at -0.01) while answering the first set of questions. S20's bias then decreased to -0.28 for the second set of revisions sets, which is consistent with what they reported during the interview.

On the other hand, students from the expandable model condition were trying to complete more activities to help align their performance with their confidence level. One student from the expandable model group claimed *'I didn't do the questions one time, I did it multiple times as I am going to the topic, I was using the self-confidence more seriously... whichever questions I was not confident I was noted down and then I will do another go until I get right'* (S18). Comparing the log data against



the qualitative evidence, S18 re-did all questions four times and stopped re-doing questions when they were all answered correctly. This student's bias score also reduced through the study process, from -0.11 the first time through the questions to -0.05 on their fourth and final round of studying.

Student use of the system to help them improve their knowledge monitoring skills may also indicate their self-regulated learning could be improved. This difference only occurred in the first set of questions and not in the second set. A Bonferroni-corrected independent t-test (with equal variance assumed) for low-achieving students in Set 2 was ( $t(15) = .19, p = 1.000, d = 0.09$ ). Those in the control condition had a mean of -0.043 and standard deviation of 0.19. Those in the expandable model condition had a mean of -0.06 and standard deviation of 0.18; therefore, their results appear to be similar. This evidence shows that low-achieving students who were assessing their confidence judgement in the first set of questions were more aware of their knowledge. Students from the expandable model claim that *'seeing your knowledge and then you get to see what the system thinks of your knowledge and then that's really helps'* (S33).

#### ***High-achieving students' confidence judgement by condition:***

A Bonferroni-corrected independent t-test (with equal variance assumed) showed no difference ( $t(12) = 1.54, p = .298, d = 0.89$ ) in the question Set 1 bias between the control group ( $M = .11, SD = .15$ ) and expandable model group ( $M = -.01, SD = .12$ ), when only high-achieving students are considered.

The over-confidence shown from the control condition could be because some high-achieving students did not take the confidence assessment seriously, for example, *'I think most of the time I will click very sure, just click that'* (S30). Student reports of this nature match the system logs. For example, S30 had an average confidence score of 3.36, which is between the two scores ("sure" and "very sure") and a bias score of 0.14, indicating that their self-confidence rating was higher than their actual ability. Similarly, another student from the control condition (S12) had an average confidence of 3.8, a bias of 0.29, and claimed that *'most of the time, I will choose very confident or just in the middle'*.

In contrast, students from the expandable model condition were paying attention to their confidence judgements with the aim of aligning their confidence with their performance or, at a minimum, they did not want to be wildly overconfident in their self-assessments. For example, S21 *'didn't want to answer very sure in case I got wrong or completely not sure'*. This student's log data shows an average confidence score of 2.96 and a bias score of 0.01, showing that self-reflection about one's abilities might occasionally induce self-doubt.

There was no significant difference between visualisation conditions for student bias when answering the Set 2 questions using a Bonferroni-corrected independent t-test with equal variance assumed:  $t(11) = 0.17$ ,  $p = 1.000$ ,  $d = .01$ . High-achieving students in both conditions had the same average bias of .00 and the same variability in their bias ( $SD = .25$ ). The bias of 0 across conditions indicates that high-achieving students had a more aligned model between their confidence and their performance when answering the questions in Set 2. This suggests that

high-achieving students were more skilled in monitoring their knowledge and were able to align their confidence with the results of their assessments rather quickly.

#### **6.4.3 Students' bias of confidence judgement over time: potential practice effects**

There was a main significant effect on participant bias from Set 1 to Set 2. A paired sample t-test was conducted to explore these changes in student bias for each condition, classified by low- and high-achieving students, and how this change could influence their performance.

##### ***Practising confidence judgement for low-achieving students:***

A Bonferroni-corrected paired sample t-test showed a large difference in bias from Set 1 to Set 2 for low-achieving students in the control condition (equal variance assumed), as shown in Table 6-3. This shows that students from the control condition had a large change in their tendency to be over-confident in Set 1 to being slightly under confident in Set 2 (see Table 6-3). Students from the expandable model condition had a stable confidence judgement while answering the questions from both Sets 1 and 2, as shown by the non-significant difference in their confidence judgements using a Bonferroni-corrected t-test (see Table 6-3).

Table 6-3. Bonferroni-corrected paired sample t-test for low-achieving students for students' bias, significant difference indicated in bold.

<b>Bonferroni-corrected paired sample t-tests</b>						
	Set 1 <i>M (SD)</i>	Set 2 <i>M (SD)</i>	<i>t</i>	<i>df</i>	<i>p-value</i>	<i>d</i>
<b>Bias</b>						
<b>Control</b>	.12 (.15)	-.04 (.19)	6.02	6	<b>.002</b>	2.27
<b>Expandable</b>	-.06 (.19)	-.06 (.18)	.23	9	1.000	.07

In addition, the corrected Wilcoxon signed rank test revealed a significant increase in control group student performance from Set 1 to Set 2 for low-achieving students (see Table 6-4). This large increase in students' performance from Set 1 to Set 2 indicates students were able to use feedback about their performance, while answering the first set of questions to reflect on their learning. Although low-achieving students' bias of confidence judgement from the expandable model was stable, they had a large increase in their performance from Set 1 to Set 2 (see Table 6-4).

Table 6-4. Bonferroni-corrected Wilcoxon signed rank test for low-achieving students for students' performance, significant difference indicated in Bold.

<b>Bonferroni-corrected Wilcoxon signed rank test</b>					
	Set 1 <i>Mdn (IQR)</i>	Set 2 <i>Mdn(IQR)</i>	<i>Z</i>	<i>p-value</i>	<i>r</i>
<b>Performance</b>					
<b>Control</b>	.52(.11)	.68(.08)	-2.384	<b>.034</b>	.90
<b>Expandable</b>	.58(.11)	.74 (.13)	-2.549	<b>.022</b>	.39

This stability in students' bias of confidence judgement indicates that those in the expandable model condition tried to align their confidence accurately with their performance, while answering both Sets 1 and 2. This was reflected by the large increase in their performance.

Low-achieving students clearly benefited from viewing model alignment, which was associated with positive changes in their performance. This change might be from students identifying their weaknesses while answering Set 1 questions, making them aware of their knowledge. This awareness could then be used to review the unknown concepts identified by OLMlets before they answered the second set of questions. Low-achieving students from the control condition claimed *'First I start going to revise whatever it was in OLMlets covered. If I got stuff wrong, I will go back and I will read all the stuff and if I got the stuff right then I will not worry about it'* (S26). They also used OLMlets *'to know how much I know'* (S20), and low-achieving students from the expandable model condition thought that OLMlets was beneficial to support their learning because *'I can understand what I know and what I didn't know'*. (S33) and it *'helped me to get the knowledge I was missing'* (S18).

***Practising confidence judgement for high-achieving students:***

For those in the control condition, there was no significant change in assessing their confidence judgement (bias) from Set 1 to Set 2 for high-achieving students (Table 6-5). Moreover, there were no significant differences in students' bias of confidence judgement for the expandable model condition from Set 1 to Set 2 (Table 6-5).

Table 6-5. Bonferroni-corrected paired sample t-test for high-achieving students for students' bias.

Bonferroni-corrected paired sample t-tests						
	Set 1 <i>M (SD)</i>	Set 2 <i>M (SD)</i>	<i>t</i>	<i>df</i>	<i>p-value</i>	<i>d</i>
<b>Bias</b>						
Control	.09 (.16)	.00 (.25)	1.92	5	.226	.78
Expandable	-.01 (.12)	.00 (.25)	.19	6	1.000	.07

Students' performance was stable from Set 1 to Set 2 in both conditions (Table 6-6). These results show that high-achieving students accurately assessed their confidence, regardless of the visualisation to which they were exposed. Table 6-6 shows the result of the Bonferroni-corrected paired sample t-test for high-achieving students' performance between the two groups.

Table 6-6. Bonferroni-corrected paired sample t-test for high-achieving students for students' performance.

Bonferroni-corrected paired sample t-tests						
	Set 1 <i>M (SD)</i>	Set 2 <i>M (SD)</i>	<i>t</i>	<i>df</i>	<i>p-value</i>	<i>d</i>
<b>Performance</b>						
Control	.69 (.05)	.74 (.11)	1.35	5	.472	.48
Expandable	.76 (.08)	.73 (.14)	.983	6	.726	.44

#### 6.4.4 System usage by condition

Accessing the learner model showed that students took responsibility for their learning, and reacted by answering more questions or reviewing material in the

area that needed improvement. This strategy was followed by both low- and high-achieving students across conditions. This strategy can be seen in Table 6-7, which shows the median number of model views for the low- and high-achieving students from both conditions. Any apparent differences in the number of times low-achieving students viewed the OLM across conditions that may be suggested by the descriptive statistics in this table are not supported by the results of a Bonferroni-corrected independent t-test with equal variance assumed for both Set 1 ( $t(16) = .592, p = 1.000, d = 0.27$ ) and Set 2 ( $t(15) = .450, p = 1.000, d = 0.22$ ). There was also no difference between the number of times the OLM was viewed for high-achieving students across conditions for Set 1 ( $t(12) = .518, p = 1.000, d = 0.29$ ) and Set 2 ( $t(11) = .314, p = 1.000, d = 0.18$ ).

*Table 6-7. Median and inter-quartile range of number of model views for both ability levels in both conditions.*

	Model views (Set 1)			Model views (Set 2)		
	<i>n</i>	<i>Mdn</i>	<i>IQR</i>	<i>n</i>	<i>Mdn</i>	<i>IQR</i>
<b>Low-achieving</b>						
Control	7	23	91.00	7	46	67.00
Expandable	11	44	63.25	10	25	49.50
<b>High-achieving</b>						
Control	7	23	28.75	6	22	28.50
Expandable	7	17	44.00	7	24	39.00

Similar to the viewing of the open learner model, no differences were detected between conditions for the number of questions answered. A Bonferroni corrected Mann-Whitney U-test showed that there was no significant difference

between low-achieving students from the control group, and the expandable model group for number of questions answered in Set 1 ( $U = 36.5, p = .854, r = .04$ ) and in Set 2 ( $U = 24.5, p = .296, r = .25$ ). There was also no significant difference between high-achieving students from the control group and the expandable model group for number of questions answered, using a Mann-Whitney U-test in Set 1 ( $U = 22.5, p = .790, r = .07$ ) and in Set 2 ( $U = 12, p = .195, r = .36$ ). Table 6-8 shows the descriptive statistics for low- and high-achieving students for number of questions answered in both groups.

*Table 6-8. Median and inter-quartile range of number of questions answered for both low-and high-achieving students for both conditions.*

	No. Questions (Set 1)			No. Questions (Set 2)		
	<i>n</i>	<i>Mdn</i>	<i>IQR</i>	<i>n</i>	<i>Mdn</i>	<i>IQR</i>
<b>Low-achieving</b>						
Control	7	75	75	7	50	75
Expandable	11	62	55.25	10	37.5	51
<b>High-achieving</b>						
Control	7	28	13.25	6	25	31.25
Expandable	7	25	32	7	47	25

Analysing student log data shows that students from each condition re-did the full question set multiple times (see Figures 6-5 and 6-6).

Figure 6-5 shows the number of visits to the entire question set for low-achieving students from the control condition. This behaviour was explained by one student who said *'If I got stuff wrong then I will go back and I will read all the stuff'* (S26). Another low-achieving student from the control condition claimed *'If it was wrong*



*I look for the answers straight away I look at the topic in the book and then do the questions again'* (S20). This quote also suggests that students revisited questions to improve their understanding of the topic, as indicated by S26 and S20 in Figure 6-5.

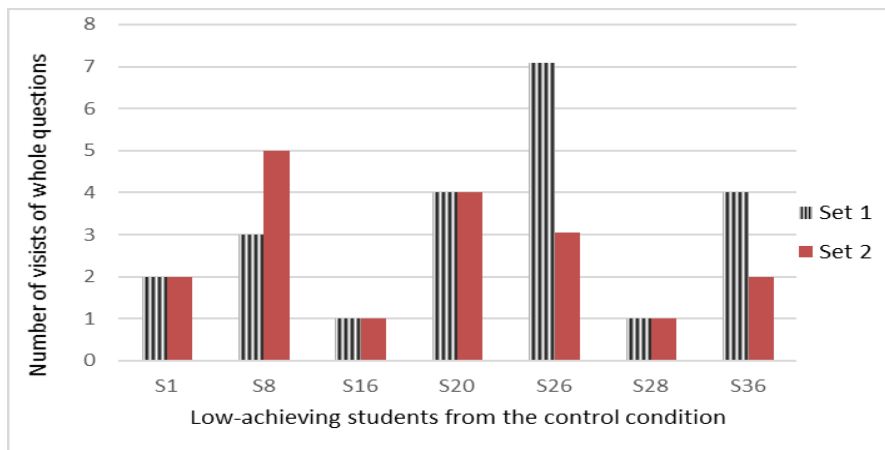


Figure 6-5. Number of times the low-achieving control-group students visit the whole question set.

Like these students, low-achieving students from the expandable model condition enacted the same strategy. For example, S9 in Figure 6-5 visited all of question Set 1 five times. The student claimed to do *'OLMlets to see where I roughly was and then normally the green bar was quite low, and then I did some revision on the kind questions asked and then I go back and do it'* (S9). Student S18 also visited the whole question set four times while answering Set 1, and five times while answering Set 2. This student believes that using OLMlets is strategic in preparation for the class test, *'I was using it to revise for tests'*.

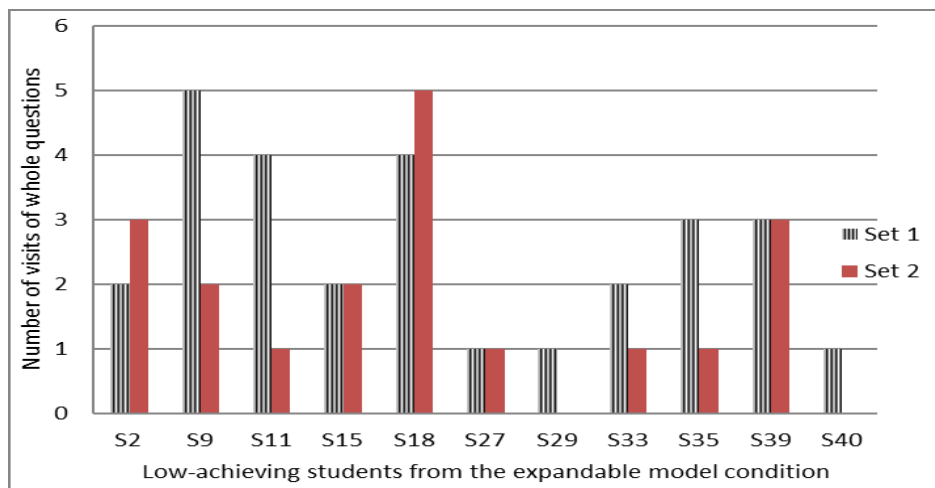


Figure 6-6. Number of times the low-achieving students from the expandable model condition visited each question set in its entirety.

With the exception of two students (S12, S24), the high achieving students who were in the control condition attempted the whole question set only once (see Figure 6-7). S12 claimed that OLMlets is a good learning resource and explained that he *'used OLMlets before I revise for one topic and then after I revised the whole booklet'*.

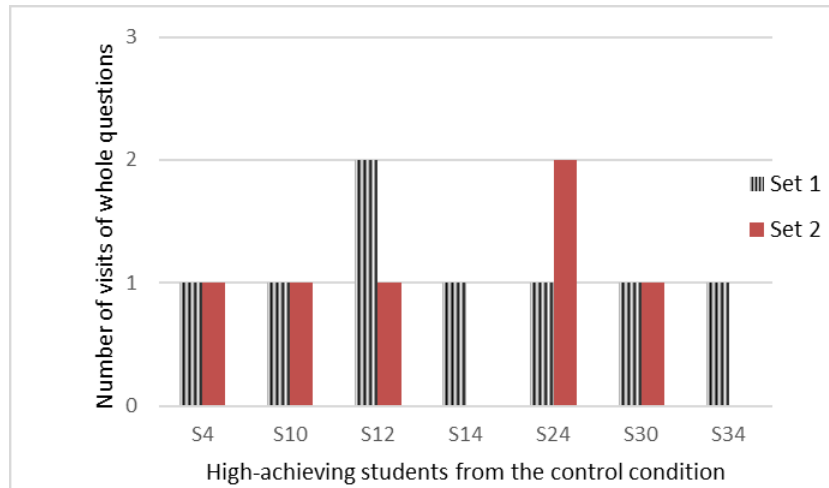
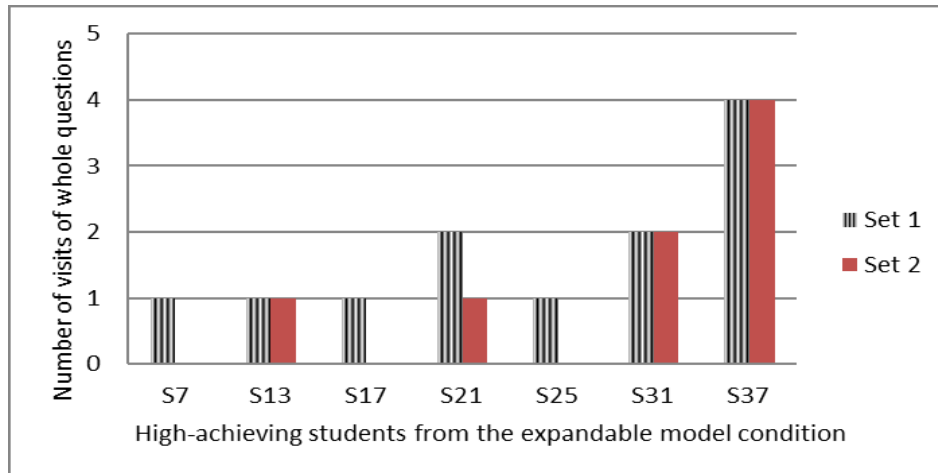


Figure 6-7. Number of times the question sets were completed for the high-achieving students from the control condition.

High-achieving students from the expandable model condition used the same strategy as their peers from the control condition. Figure 6-8 shows that students were motivated to answer the first set of questions, but not all students were motivated to answer all of the questions in Set 2. One of these students (S37) answered both question-sets four times, and stated that, “*doing OLMlets really helped with like gaining new knowledge and consolidating the old knowledge*”. This student also thought that OLMlets supported exam preparation because it allowed him to “*practise for the class test*”.



*Figure 6-8. Number of visits to question sets for high-achieving students from the expandable model condition.*

Comparing the question of completion behaviour of low-achieving students across conditions (Figures 6-5 and 6-6) to that of high-achieving students across conditions (Figure 6-7, Figure 6-8), may indicate that the OLM motivated low-achieving students to perform learning activities within the system. This claim is supported by the significant difference between low and high-achieving students, regardless of their condition, in number of attempts for the whole questions set using Mann-Whitney U-test for Set 1 ( $U = 56.5, p = .007, r = .47$ ) with a mean rank of 20.36 for low-achieving students and 11.54 for high-achieving students. This suggests that low-achieving students tried to do more activities to prepare for their actual class test, whereas high-achieving students might have felt confident about their preparation for their test. The significant difference in number of attempts of all questions was not observed between low- and high-achieving students in Set 2 using Mann-Whitney U-test ( $U = 102, p = .673, r = .07$ ) with a mean rank of 14.88 for low-achieving students and 16.21 for high-achieving students).

## **6.5 Discussion**

This study examined students' bias in their confidence judgement while students were interacting with the OLM in the OLMlets system. In this section, the implications of the results for supporting students' monitoring tasks as they regulate their learning are discussed.

### **6.5.1 The effect of OLMlets and visualising alignment on students' bias of confidence judgement**

In this study, the findings show that visualising alignment supported students to align their confidence to their performance, resulting in far more accurate confidence judgements ( $d = 1.02$ ), regardless of their ability level. This effect was not found for those in the control condition, where students used standard skill meters. This result suggests that providing students with separate information about their confidence and their performance makes students more aware of the alignment between their knowledge and their confidence, thus supporting the development of their knowledge monitoring skills. The result is consistent with prior research, where students who received feedback about their performance and their confidence had more accurate confidence judgement than those who only received feedback about their performance (Renner and Renner, 2001; Callender et al., 2016).

Exploring the influence of visualising alignment based on student performance showed that low-achieving students from the expandable model condition sought to align their confidence with their performance, which led to their confidence being more stable across question sets and a large increase ( $d = 1.45$ ) in their

performance from Set 1 to Set 2.

The large improvement ( $d = 2.27$ ) in the accuracy of the control group's low-achieving students' bias of confidence judgements from question Set 1 to Set 2 suggests that the simple act of practicing self-assessment can be beneficial. This self-assessment demonstrates reflection in action, which, as expected, positively impacted on their performance (Schon, 1983). This may provide internal feedback that supports student monitoring of their knowledge. This improvement in students' confidence judgement was also accompanied by a large increase ( $r = .9$ ) in their learning performance. This result is in line with the theory suggested by Koriat (1997), that practising confidence judgement may improve students' ability in monitoring their knowledge. The result also confirms prior findings that practising confidence judgement and using internal feedback can enhance self-regulated learning (Butler and Winne, 1995) in a new instructional domain. The findings here are consistent with research where overconfidence reduces when students assess their confidence level after each question and receive feedback about their performance (Renner and Renner, 2001).

Visualising alignment in OLMlets was associated with more accurate confidence judgements. This study showed the importance of visualising alignment to support students' confidence judgement and how this support benefits students' learning. Prior findings show that OLMs improved students' self-assessment accuracy (Mitrovic and Martin, 2007; Kerly and Bull, 2008). Additional evidence was provided of this effect at a time when the need for the replication of previous findings across contexts was growing. These findings expand on previous work

by implying that adding information to the OLM about how student beliefs and performance are aligned can raise student awareness of specific aspects of their metacognitive skills (i.e., knowledge monitoring and self-assessment).

The findings of both studies in this chapter and Chapter 5 answer the thesis research question: visualising alignment can support students' confidence judgement. The findings show that students who can view the alignment information in their learner model were more accurate in judging their confidence while answering the system-automated assessments. This finding supports the theory of this thesis that a student's experience of misalignment between their confidence and performance could motivate them to work hard and resolve the misalignment. Based on the students' own perceptions from the interview data, students who could view the alignment between their confidence and their performance were more motivated to align both their confidence and their performance. As a result, they were more serious in their self-assessment skills, which was supported by the evidence of the findings from both the quantitative and the qualitative methods used in this study.

The previous study showed that adding the alignment information to students can have an impact on their confidence, where students who can view the alignment experienced gain in their confidence compared to the control condition, who could not view the alignment information and had a decrease in their confidence. Both studies suggest that incorporating students' confidence in their knowledge and the system assessment of their knowledge into an OLM could benefit students

self-monitoring skills, particularly for low-achieving students and students' confidence overall.

This thesis presents different methods that could be used for showing the alignment between students' confidence and performance (discussed in Chapter 3). The next study in Chapter 7 investigates the impact of the representation of the alignment when using either a text-based or a graphical learner model integrated to the OLMlets system. The survey discussed in Chapter 4 showed that students preferred the use of the opacity of the colour as their preferred method of visual variable to indicate this information (i.e. alignment) within the OLM. The graphical interface uses the same expandable model used in this study, which indicates the alignment using the opacity of the colour. The findings of this study will help provide guidelines for system designers, revealing whether the representation of the alignment is important to have more accurate confidence judgement.



# **Chapter 7 : EVALUATING THE EFFECT OF THE REPRESENTATION OF ALIGNMENT ON STUDENTS' CONFIDENCE JUDGEMENT**

---

## **7.1 Overview**

The study presented in Chapter 6 highlighted the importance of visualising alignment within the OLMlets OLM to support students' confidence judgement and reflection, particularly for low-achieving students. In this chapter, the study will expand upon that work. The same students participated in this study as that in Chapter 6; however, here they are exposed to different visual representations of model alignment. Prior research has argued that the presentation method (i.e., visual or textual) may influence students' metacognitive judgement (Burkett and Azevedo, 2012). Thus, in this study, the expandable model treatment condition from the previous study is retained, and a text-based version of this model alignment approach is provided at the second condition. The study used a mixed methods approach. Qualitative data was collected through semi-structured interviews conducted after both studies had been completed (as described in Chapter 6). These interviews explored students' perceptions of their experiences in using the system across both studies, to understand their use and their reactions to the

different methods for representing their learner model better. Quantitative data was collected through system logs while students interacted with the OLMlets system. As such, this follow up study hypothesised that bias of students' confidence judgement will differ according to the mode used to present alignment information (text vs. graphical). This study addresses the third research question of this thesis: 'How is information presentation format important in terms of adding the alignment information in an OLM?'

## **7.2 Introduction**

Learner models in Intelligent Tutoring Systems (ITSs) are often externalised for students through visualisation, at which point they are called OLMs.

Prior research on ITSs has used text (Bull and Pain, 1995; Kay, Li and Fekete, 2007) or visual display (Kerly and Bull, 2008; Zapata-Rivera and Greer, 2004) to provide feedback about students' confidence in the correctness of their answers and their performance, based on computer evaluation. However, these studies did not investigate the impact of the alignment representation on either students' bias of confidence judgement or students' behaviour in using the system using a control study.

This chapter presents a study that investigates the importance of the representation of the alignment. This study tested whether providing different representations of the learner model that shows this alignment (skill meter, text-based) can influence students' bias of confidence judgement and their behaviour in using the system. Although the visual-based approach could display more information from the student model, both representations follow the same modelling process

discussed earlier in Chapter 3. The implications of this study can further guide the system designer to which method to use when considering students' confidence judgement in the correctness of the answers given in OLMs.

The study is a follow-up of research presented in the previous chapter. Students' bias of confidence judgement indicates whether students' tendency in judging their knowledge is either over- or under-confident. The bias index 0 indicates that the students aligned their confidence from their own self-confidence assessment that the answers are correct to their performance based on the computer evaluation. The study compared two versions of the learner model that include the alignment information text-based learner model and graphical learner model (skill meter). The hypothesis of this study was as follows:

**H1:** Bias in student confidence judgement will differ according to the mode used to present alignment information (text vs. graphical).

The results from this study can help system designers in OLMs and ITSs, in terms of how to present the alignment information to support metacognitive skills.

## **7.3 Method**

### **7.3.1 Participants**

As already stated, participants in this study were the same as for the study presented in Chapter 6. Students were in the second-year of an undergraduate program in the School of Electronic, Electrical and Systems Engineering, University of Birmingham in the United Kingdom. There were 46 students in the class. Of those 46 students, 38 students had valid logs (number of questions > 10 and

number of times viewed the model >2). Use of the system was voluntary and students used it in their own time. Students were identified anonymously using identifier numbers from S1 to S46. The student identifier number indicates the same students across both studies in both Chapters 6 & 7.

Students were classified into two levels, based on the median score of their performance of their first attempt at solving the first set of questions of the system's automated assessment. In this study, the median score was higher (*Mdn* = 76%) than the median score for the previous study (*Mdn* = 62%). Students who were below the median (*Mdn* = 76%) classified as low-achieving students and those above or at the median score classified as high-achieving students. The number of participants in each condition is shown in Table 7-1.

*Table 7-1. Number of participants in each condition classified by the median score of students' performances in Set 1.*

<b>Ability level</b>	Median	<b>Visualisation Conditions</b>	
		Expandable	Text-based
Low-achieving	<76%	11	7
High-achieving	>=76%	10	10
<b>Total</b>		21	17

### **7.3.2 Study design**

The study took place in the second semester and was two weeks in duration, and students volunteered to use the system as a revision tool to prepare for their final exam. Students were enrolled in a computer science course called *Unified*

*Modelling language-UML* that was integrated into OLMlets to fulfil the study aims. Two sets of review questions (covering six lectures of the course curriculum) were integrated into OLMlets. Each set of questions included 25 multiple-choice questions. Set 2 was uploaded to OLMlets one week after uploading Set 1. This was done to explore students' practice effect to judge their confidence while answering the system's automated assessment.

There were two visualisation conditions. The conditions were based on how the alignment information was represented (textual vs. graphical) and students were classified according to their ability level, as either low- or high-achieving. Because the participants for this study were the same as for the previous study (Chapter 6), they were switched between conditions so that they would experience different OLM versions. The graphical representation (expandable model) was the same as that used in the previous study. Figures 7-1 and 7-2 show the learner model for one of the students who participated in the study, with different representations of model alignment. This enables the reader to compare between both views (the expandable model in Figure 7-1 and the text-based model in Figure 7-2). To explore students' perceptions about their experience of different versions of their learner model representation, semi-structured interviews were conducted at the end of the study.

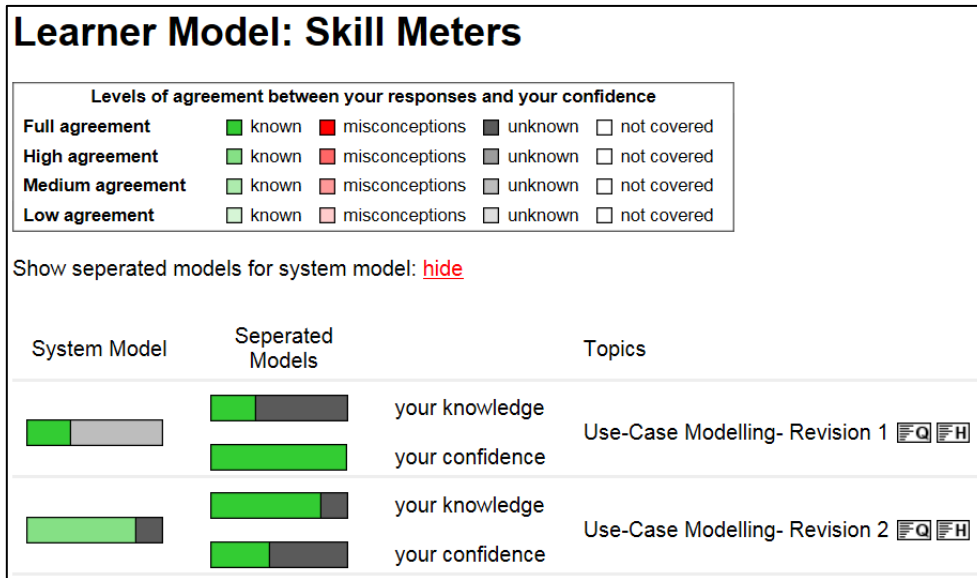


Figure 7-1. The expandable model shown in the expanded view specified in this study by student S25.

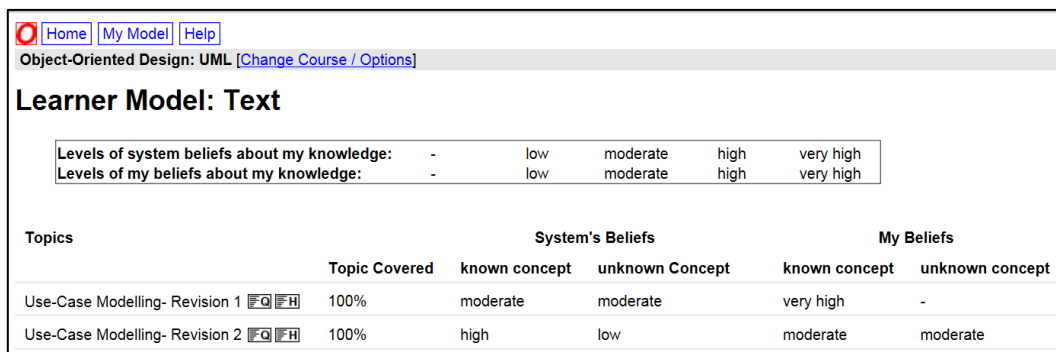


Figure 7-2. A text-based learner model within OLMlets shows both (System's Beliefs, My Beliefs) specified in this study by student S25.

### System Training

A 15-minute presentation was given to participating students during class time, to provide an introduction on how to use the system and explain its features. All students received the same presentation time, so each group was able to see the OLMlets study condition. Students were given a user manual with a visual, step-

by-step screenshot guide to help students understand the system. Each student had a username and password to log in to the system.

### **7.3.3 Design of the questions integrated to OLMlets**

The study was based on a programming course (*Unified Modelling Language-UML*). Two sets of review questions focused on the topic 'Use-case Modelling' from the UML course curriculum. In each of Set 1 and Set 2, there were 25 multiple-choice questions. Questions integrated into OLMlets were classified according to Bloom's Taxonomy (Bloom, 1956) (explained in chapter 6) and the course instructor rated the questions on a scale from 1 to 5 (1 being easy and 5 most difficult). The instructor rating of the difficulty of questions was performed independently of the Bloom categorisation. The difficulty rating of each Bloom category was the average score of the level of difficulty of questions rated by the instructor that belong to the same Bloom category. For example, the average difficulty rating for the 14 questions included in the Bloom category 'knowledge' (Table 7-2) was 2, based on the instructor's viewpoint. Using two methods of categorising the questions ensured that both sets of revisions (Set 1 and Set 2) did not differ in level of difficulty or type of questions; thus, they did not influence students' bias of confidence judgement between both sets. The two sets were used to enable comparison between students' bias of confidence judgement after practising their confidence judgement in Set 1. The questions were categorised according to Bloom's Taxonomy, and questions mainly focused on students' knowledge and comprehension, with a small number of questions concentrated on how students apply new concepts. The top level of Bloom considered to be

outside the scope of the multiple-choice assessments. Table 7-2 shows the percentage of questions and difficulty of the questions based on the instructor's viewpoint.

*Table 7-2. Percentage of questions, difficulty rating based on instructor's opinion and the percentage of correct responses of student answers for each Bloom Taxonomy category.*

	Knowledge	Comprehension	Application	Analysis	Synthesis	Evaluation
<b>Set 1</b>						
Number of questions (N =25)	14	10	1	0	0	0
Average difficulty rate (1 → 5)	2	3	3	-	-	-
<b>Set 2</b>						
Number of questions (N = 25)	14	7	4	0	0	0
Average difficulty rate (1 → 5)	3	3	2	-	-	-

Here are some examples of some of the questions integrated to OLMlets based on Bloom's Taxonomy.

**Knowledge:**

In UML diagrams, relationship between object and component parts is represented by

- a) Ordination.
- b) Aggregation.
- c) Segregation.
- d) Increment.

**Comprehension:**

All of which are valid relationship in Use Case Diagrams except

- a) Generalisation



- b) Include
- c) Extend
- d) Exclude

**Application:**

If 'Submit loan request' and 'Offer line of credit' both uses 'Perform credit check' use case, the relationship between Perform credit check and other used cases is:

- a) Extend relationship.
- b) Include relationship.
- c) Generalisation relationship.
- d) No relationship.

**7.3.4 Data collection**

The study used both quantitative and qualitative methods to collect the data. Quantitative data consisted of system logs of the students' interaction with the system, stored in a database, and included performance data on learning tasks and their confidence rating for each question answered. Semi-structured interviews provided the qualitative data used to explain students' results (from the quantitative data), to understand students' perceptions of using the system, and to explore their experiences of using different OLM visualisations after both studies had ended. Quotations were used to augment the quantitative results.

**7.3.5 Measures**

The measurement in this study was based on system logs of the first attempt of answering the questions (N = 25). Students' confidence was derived from system logs when students assessing their confidence after each question answered the system-automated assessment. Students' performance was calculated based on whether the answer was or was not correct. The confidence level was calculated

based on the average of the confidence scores reported by students, after answering the questions in the OLMlets system. Student's bias score was calculated based on the first visit (for the questions and for both revision sets, bias 1, bias 2), see Equation [5]. The sign magnitude for the bias score indicates whether the student was under- or over-confident. A negative score indicates under-confidence, whilst a positive score indicates over-confidence, and a score of 0 indicates that students' performance was aligned with their confidence judgement. (Schraw, 2009).

$$\text{Bias} = \frac{\sum_1^n (\text{confidence} - \text{performance})}{n} \quad [5]$$

### 7.3.6 Quantitative analysis

Two independent variables were used (visualisation condition and students' ability level). Two conditions were used (expandable model, text-based model). Students' ability level was classified according to two levels: low-achieving, or high-achieving. Four dependent variables were measured. The first dependent variable was the students' bias of confidence judgement (bias 1 and bias 2). The second dependent variable was students' performance (performance 1 and performance 2), based on their performance in providing correct answers according to the computer evaluation. Two dependent variables were used to explore students' behaviour is using the system: number of questions answered, and number of times the model was viewed.

A Shapiro-Wilk test was used to assess the normality of data. The Mann-Whitney U-test was used to identify any significant differences between the groups for

non-parametric data. A Wilcoxon signed rank test was used to identify any significant difference within subjects for non-parametric data. A paired sample t-test was used to identify significant differences within subjects for normal distributed data.

The effect size ( $r$ ) reported in this study for non-parametric tests is based on Cohen's guidelines for effect size (Cohen, 1988). An effect size of .01 indicates a small effect, 0.3 a medium effect, and .5 a large effect. The effect size of Cohen's  $d$  was calculated for t-tests and followed the guidelines of 0.2 as small effect, 0.5 for medium effect, and 0.8 for large effect (Cohen, 1988).

### **7.3.7 Qualitative analysis**

The qualitative data were collected through semi-structured interviews after both studies were completed (Chapters 6 and 7). Appendix 13 provides the questions asked in the interview. Students' interviews were recorded and transcribed into papers. Quotations were then classified into themes and used to augment the quantitative data. Appendix 14 presents transcriptions classified into themes, such as exam preparation, understanding a topic, revision strategy, confidence in understanding, and visualisation preference.

## **7.4 Results**

### **7.4.1 Normality of data**

A Shapiro-Wilk test showed student bias was normally distributed ( $p > .05$ ) in most cases. The two exceptions to this were the bias 1 measure for high-achieving students in the expandable model condition ( $p = .031$ ), and the bias 2

measure for low-achieving students in the expandable model condition ( $p = .003$ ). Performance in both Sets 1 and 2 was normally distributed, except for high-achieving students from the text-based condition ( $p = .033$ ). The number of times students viewed the model was normally distributed for Set 1 ( $p > .05$ ), and not normally distributed for Set 2 ( $p < .05$ ). The number of questions students answered was not normally distributed ( $p < .05$ ). Appendix 15 shows students logs data for all students for the data used in this study.

#### 7.4.2 Students' confidence judgement: visual vs. textual OLMs

Students in both conditions (text-based model, expandable model) could compare their performance to their confidence. Table 7-3 presents the descriptive statistics for student bias of confidence judgement and their performance in Set 1 and Set 2 assessments.

*Table 7-3. Students bias for low- and high- achieving students in Set 1 and Set 2 for both conditions (text-based learner model, expandable learner model).*

	<i>n</i>	<i>Bias 1</i>	Set 1 (Performance.)	<i>n</i>	<i>Bias 2</i>	Set 2 (Performance.)
		<i>M (SD)</i>	<i>M (SD)</i>		<i>M (SD)</i>	<i>M (SD)</i>
<b>Low-achieving</b>						
Text-based	7	.1(.18)	.58 (.08)	4	.03 (.30)	.75 (.22)
Expandable	11	.13 (.26)	.56 (.14)	10	.03 (.28)	.69 (.13)
<b>High-achieving</b>						
Text-based	10	-.16 (.25)	.82 (.08)	7	-0.5 (.23)	.79 (.23)
Expandable	10	-.11 (.34)	.84 (.08)	10	.02 (.28)	.81 (.14)

The results showed that there was no significant difference in students' bias of confidence judgement accuracy between both conditions for low-achieving students, using a Bonferroni-corrected Mann-Whitney U-test for Set 1 ( $U = 38$ ,  $p = 1.000$ ,  $r = .01$ ) with a mean rank 9.43 for the text-based model and 9.55 for the expandable model. Further, there was no significant difference for low-achieving students between both conditions in Set 2 using a Bonferroni-corrected Mann-Whitney U-test ( $U = 17$ ,  $p = 1.000$ ,  $r = .11$ ) with a mean rank of 6.75 for the text-based OLM and 7.80 for the expandable model.

There was no significant difference between conditions for high-achieving students using a Bonferroni-corrected Mann-Whitney U-test for Set 1 ( $U = 40$ ,  $p = .900$ ,  $r = .17$ ) with a mean rank 9.50 for the text-based and 11.50 for the expandable model. There was also no significant difference between both conditions for high-achieving students in Set 2, using a Bonferroni-corrected Mann-Whitney U-test ( $U = 32$ ,  $p = 1.000$ ,  $r = .07$ ) with a mean rank of 8.57 for the text-based and 9.30 for the expandable model OLM.

Comparing students' bias of confidence judgement between low- and high-achieving students in each condition, there was a difference between the low- and high-achieving students who were using the text-based OLM for Set 1. This difference was observed for bias 1 using a Mann-Whitney U-test ( $U = 12$ ,  $p = .025$ ,  $r = .54$ ) with a mean rank of 7 for low-achieving students and 10 for high-achieving students. However, a difference was not observed between the low- and high-achieving students who used the expandable model, using a Mann-

Whitney U test for bias 1 ( $U = 35, p = .159, r = 30$ ) with a mean rank of 11 for low-achieving students and 10 for high-achieving students.

Although the difference in students' bias of confidence judgement was observed between low-and high-achieving students for the text-based condition in Set 1, it was not observed in Set 2 using a Mann-Whitney U-test ( $U = 11.5, p = .636, r = .14$ ) with a mean rank of 6.63 for low-achieving students and 5.64 for high-achieving students. Further, there was no significant difference between low- and high-achieving students for the expandable model condition in Set 2, using a Mann Whitney U-test ( $U = 46, p = .762, r = .06$ ) with a mean rank of 10.9 for low-achieving students and 10.1 for the high-achieving students.

Table 7-4 presents the statistical analysis, using a Bonferroni-corrected Wilcoxon signed rank test to indicate if there was a significant increase in students' performance from Set 1 to Set 2, based on the computer evaluation that the answers given are correct for both students' ability level in each condition.

*Table 7-4. Bonferroni-corrected Wilcoxon Signed Ranks Test for students performance from Set 1 to Set 2, Significant increase from Set 1 to Set 2 is indicated as bold and italic.*

Ability Level	Condition	Z	P
<b>Low-achieving students</b>			
	Text-based	-.730	.93
	Expandable	-2.504	<b><i>0.24</i></b>
<b>High-achieving students</b>			
	Text-based	-.169	1.00
	Expandable	-1.126	.52

From Table 7-4, it is evident that low-achieving students from the expandable model condition had a significant increase in performance from Set 1 to Set 2. This significant increase was not observed in low-achieving students from the text-based condition. This may be because students from the expandable model benefited from the graphical representation, and the effect of prior experience from the previous study (Chapter 6) for the expandable model condition had an effect on how students could interpret their learner model. In contrast, students from the text-based condition were experiencing the text-based representation for the first time. This may also indicate that students using the text-based model may have had difficulty interpreting the text-based version of the OLM, or were not interested in interacting with text-based information. As stated by one student from the text-based OLM condition, *'the text-based you do not really engage it well, I think a mixture of both (visual and text) is something that I will prefer'* (S33).

High-achieving students from both conditions had no significant improvement in their performance from Set 1 to Set 2. One reason for this might be that the scores of high-achieving students are such that there was no room for improvement.

### 7.4.3 System usage

Table 7-5 presents the descriptive statistics for the number of questions answered for both low- and high-achieving students in both conditions.

*Table 7-5. Mean and standard deviation for low-and high-achieving students in both conditions.*

	Number of questions answered (Set 1)			Number of questions answered (Set 2)		
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
<b>Text-based</b>						
Low-achieving	7	54	32.8	4	19	14
High-achieving	10	49	41	7	38	27.8
<b>Expandable</b>						
Low-achieving	11	71	48.8	10	52	32
High-achieving	10	40	20	10	35	17.6

Table 7-6 presents the descriptive statistics for the number of times students viewed the model in both Sets 1 and 2 for low- and high-achieving students in both conditions.

*Table 7-6. Mean and standard deviation for number of times the model was viewed for both low-and high-achieving students in both conditions.*

	Model viewed (Set 1)			Model viewed (Set 2)		
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
<b>Text-based</b>						
Low-achieving	7	35	40.5	4	8	14
High-achieving	10	34	35	7	33	37.5
<b>Expandable</b>						
Low-achieving	11	59	47.3	10	46	30.8
High-achieving	10	28	24.4	10	23	20



There was no significant difference for low-achieving students in the number of questions answered using a Bonferroni-corrected Mann Whitney U-test for Set 1 ( $U = 30$ ,  $p = .878$ ,  $r = .18$ ) with a mean rank of 8.29 for the low-achieving students in the text-based model and 10.27 for low-achieving students in the expandable model. A Bonferroni-corrected Mann-Whitney U-test ( $U = 7.00$ ,  $p = 0.052$ ,  $r = 0.58$ ) showed a trend of low-achieving students from the expandable model attempting more questions in Set 2, with a mean rank of 9.8, than those who used the text-based model, with a mean rank of 4.4. This indicates that students using the graphical-based model were more motivated to continue answer more questions in Set 2 than low-achieving students from the text-based model.

There was no significant difference in the number of questions answered for high-achieving students across conditions (text-based model and expandable model) for Set 1:  $U = 40$ ,  $p = .856$ ,  $r = .18$ ; with a mean rank of 9.5 for high-achieving students from the text-based model, and 11.5 for high-achieving students from the expandable model. The same holds for Set 2; no differences were detected ( $U = 33$ ,  $p = 1.000$ ,  $r = .05$ ) where those in the text-based model answered a similar number of questions, with a mean rank of 8.7 for high-achieving students from the text-based condition as those from the expandable model condition, with a mean rank of 9.2.

There was a significant difference in the number of times the OLM was viewed for low-achieving students between the two conditions (text-based and expandable model) using a Bonferroni-corrected Mann-Whitney U-test in Set 2 ( $U = 4.00$ ,  $p = 0.020$ ,  $r = 0.67$ ). This difference in low-achieving students of number of times

the model was viewed only occurred in Set 2; it did not occur during Set 1, using a Bonferroni-corrected Mann-Whitney U-test ( $U = 20, p = 1.000, r = .39$ ). High-achieving students had no significant difference in their behaviour in viewing the model in both Set 1 and Set 2 using a Mann-Whitney U-test ( $U = 45.5, p = 1.000, r = .07$ ) for Set 1 and ( $U = 31, p = .694, r = .09$ ).

Figure 7-3 illustrates the number of visits for the whole question set ( $N = 25$ ) for the low-achieving students from the expandable model condition for both Set 1 and Set 2. In the most extreme case (S37), the student was motivated to answer the question set as many as 6 times for Set 1 and 4 times for Set 2. There was only one student (S35), who did not complete all 25 questions from Set 2.

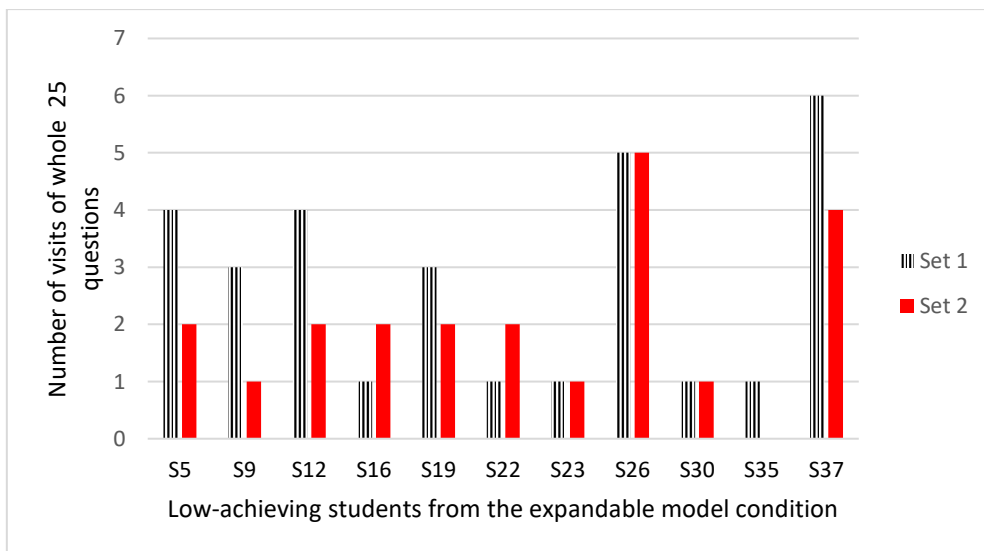


Figure 7-3. Number of visits of questions for low-achieving students from the expandable model condition.

Figure 7-4 shows the number of times low-achieving students completed each of the question sets for the text-based OLM. There were only four students who answered questions in Set 2 from the text-based condition; only three of these

students completed the entire question set. Student S7 answered only 13 questions from the 25 questions provided. This difference in system usage for low-achieving students may show that the visual display provides ease of comprehension with less time than a text-based OLM, which needs some time to determine the alignment between models through the comparison of labels. The longer time in interpreting the text-based OLM may demotivate low-achieving students from interacting with the system. A high-achieving student from the text-based condition who had used the expandable model in the first study claims '*checking my performance was easier to use in the first model*' (S21).

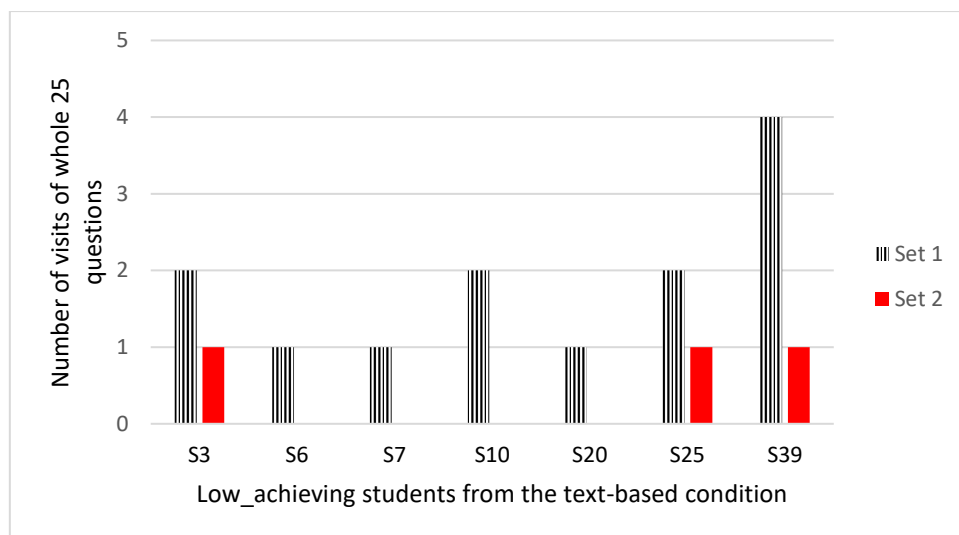


Figure 7-4. Number of visits of whole questions for low-achieving students from the text-based model condition.

Transparency as an indication of monitoring skills appears to have motivated students to aim for the opaque green colour by answering more questions, especially for low-achieving students who need more motivation to support their self-

regulation skills (VanZile-Tamsen and Livingston, 1999). A low-achieving student from the expandable model condition claimed *'if it is more transparent then I will go over the questions a bit more until I found confident'* (S30).

It might be that students' experience of the representation in the first study increased their familiarity with the skill meter representation, which may be linked to students' motivation when the graphical representation of the OLM was replaced with text. This interpretation is supported by students' perceptions. For example, one student said, *'because I got the text-based one so I was a bit of taking more time to get the right answer...I personally prefer the first one [expandable model]'* (S18). Another student who had used the standard skill meter (control condition in Chapter 6 study) preferred the standard skill meter to the text-based version from this study, *'going from high to moderate obviously, the last question's wrong...I think I prefer the skill meter'* (S20).

Students in the expandable model condition can see the two models separated by expanding the model, whereas in the text-based model the model is always expanded. Appendix 15 includes student data for all students in both conditions, including the number of times students expanded the model for students in the expandable model condition.

## **7.5 Discussion**

In this section, the discussion of the results and their implications for supporting students monitoring tasks as they regulate their learning are discussed.

### **7.5.1 Effect of the OLM representation on students' bias of confidence judgement**

This study investigated how the accuracy of students' confidence assessment was influenced by the manner in which the alignment of the two data sources was shown to learners. Essentially, this study explores whether students responded differently to a text-based version of the model than to a graphical version.

The study found a large ( $r = .54$ ) difference in bias between low- and high-achieving students who used the text-based learner model. Low-achieving students were over-confident, whereas high-achieving students were under-confident. Based on students' perceptions of their experience using the text-based OLM, students had some difficulty in observing the small change that can occur in their level of knowledge, as the text-based model did not appear to change as rapidly as the visual version. This finding confirms prior research that low-achieving students are less accurate in monitoring their performance compared to high-achieving students, who are more realistic but more likely to under-estimate their knowledge (Hacker et al., 2000). Prior research has also shown that over-confidence may lead to demotivation, which may be reflected negatively in learning (Koriat, 1997; Kelemen, 2000). This over-confidence was observed in low-achieving students in the text-based condition, which may have affected their motivation to complete question Set 2.

Data also showed a trend that students using the expandable model were more willing and possibly more motivated to interact with the system than those using the text-based version of the model. Students who had been assigned to the

graphical OLM condition answered moderately ( $r = .58$ ) more questions and viewed their OLM more ( $r = .67$ ). These behavioural tendencies may indicate that students preferred the graphical representation (i.e., skill meter), which may have benefitted student motivation more than a textual representation. Alternatively, students found that the skill meter was more familiar to them than the text-based version, which may have increased their comfort in using that graphical representation of the underlying model. This finding is in line with other studies, which indicated a student preference for using skill meters over other visualisations (Bull et al., 2014; Duan, Mitrovic and Churcher, 2010). Further, prior research has demonstrated the effectiveness of the skill meter visualisation for supporting student learning, particularly for low-achieving students (Mitrovic and Martin, 2007; Long and Alevan, 2013b).

Here, the results imply the graphical, skill meter representation of the learner model benefitted low-achieving students, with those who used the text-based version needing more information, or information that was easier to interpret.

The theory supporting this thesis explains that providing students with the information of the alignment about their performance and their confidence judgement could motivate students to reveal misalignment. These motivations were observed in the previous findings of the two studies (discussed in Chapters 5 and 6), where students who visualised alignment in their learner model were more accurate in assessing their confidence judgement, because students were more tied in to their self-assessment. In this chapter, findings show the importance of communicating the alignment between the computer's evaluation of student

performance and student confidence in the correctness of their answers as a means to support metacognitive skills and reflect on their learning.

# Chapter 8 : CONCLUSIONS AND FUTURE WORK

---

This thesis explored the alignment between students' confidence in the correctness of their answers (which would be self-assessed without external influences) with students' performance based on the computer evaluation, visualised using an Open Learner Model (OLM). The thesis visualised alignment using one of the methods proposed by Demmans Epp and Bull (2015), to indicate uncertainty or misalignment in OLMs using visual variables. The exploration of the impact of visualising alignment on students' confidence judgement in OLMs in a controlled condition has not been widely researched. This thesis described the impact of visualising alignment in natural class settings, which could aid researchers and system designers to consider the importance of visualising alignment, when supporting knowledge-monitoring skills is one of main aims of their systems.

This chapter considers the context of the research, and discusses the findings of the thesis, its contribution to the field of OLMs, and the limitation and future work in this particular area.

## 8.1 Context

This thesis reports on studies that evaluated an OLM in the context of undergraduate programming courses (*Java Programming* and *Unified Modelling Language*) and a mathematics course (*Introduction to Complex Variables*).



Participants used the system on a voluntary basis, and tended to use it as a revision tool to prepare for their class tests. Voluntary use created some difficulties and limitations that are discussed later in this chapter. The system (OLMlets) is an educational technology that models students' knowledge, based on students' answers to multiple-choice questions. OLMlets then shows aspects of this model to students, so that they can monitor their learning. OLMlets has been extended to incorporate students' confidence in the correctness of their answers, and a computer evaluation of student performance through visualisation.

## **8.2 Findings**

This thesis focuses on studies conducted in real class settings, to investigate how to support learner knowledge monitoring skills in OLMs, through visualising alignment of the computer's evaluation of students' performance and students' own perception of their performance.

This thesis aimed to answer three research questions, the first of which is 'How does visualising alignment support students' confidence judgement?' Researchers argue that providing a simple OLM could not be an effective way to support students' knowledge monitoring skills, and there should be additional scaffolding in order for the tool to be used effectively (Long and Alevan 2013c). This thesis showed how a simple OLM was used effectively to increase students' confidence and to support students' knowledge monitoring skills when the alignment information was visualised. This result confirms the theoretical approach of the benefit of OLMs as a means to support students' metacognitive skills (Bull and Kay, 2013). The experimental study in Chapter 5 illustrates how students in the

treatment condition (expandable model) who could view the alignment information in their learner model had a significant increase in their confidence gain using a paper-based pre- and post-test after OLMlets use. However, students from the control condition experienced a reduction in their confidence, indicating that the standard skill meter (where the alignment information is hidden) was not as effective as the expandable model condition. Also, my research found that students from the expandable model interacted with the system significantly more than the control condition, who were unable to view the alignment in their model. Thus, visualising alignment showed positive influences on students' confidence, which in turn had an impact on their behaviour in using the system, by encouraging them to answer more questions and view the model more. This result shows that visualising alignment supports the theory of Kulhavy and Stock (1989), that providing students with information about the discrepancy between their confidence in their performance and their performance on assessment could motivate students to work harder to resolve the misalignment.

Whilst Chapter 5 investigated the impact of visualising alignment based on paper pre- and post-test (before and after OLMlets use), the experimental study in Chapter 6 investigated the impact of visualising alignment while students interacted with the system. The study provided promising results that students from the expandable model (treatment condition) sought to align their confidence with their performance, which led students to accurately assess their confidence judgement and had a significant increase on students' performance, particularly for low-achieving students. These findings support the theory of 'experiential

learning (Dewey, 1938), where students benefited from visualising alignment, seeking to align their confidence to their performance. This result is in line with prior research, where students who received feedback about their confidence and their performance on task had more accurate confidence judgement than those who only received feedback about their performance (Renner and Renner, 2001; Callender et al., 2016).

The second research question addressed by this thesis was ‘What is the preferred method to visualise alignment for Open Learner Model users?’ The research conducted here builds on that of Demmans Epp and Bull (2015), where different methods to indicate inconsistent data in the context of OLMs were proposed.

However, exploring students’ preferred method for visualising misalignment or any other inconsistent data in the model has not been explored in the field of OLMs. This thesis fills this gap, by conducting a survey-based analysis of paired comparison questionnaire items for six visual variables applied to a skill meter visualisation. The skill meter visualisation was used as the base representation, because the skill meter was the most preferred visualisation when other visualisations were available (Bull et al., 2014; Duan, Mitrovic and Churcher, 2010). Researchers argue that users may be more motivated to use their preferred method and more able to understand the conveyed information of the visual display (Gerharz and Pebesma, 2009). Results showed that opacity of the colour was the most preferred visual variable (compared with other visual methods such as size, orientation, numerosity, added marks, and arrangement). Students

tended to prefer non organised line (i.e. arrangement) less; this result is consistent with prior research, where participants preferred more organised visualisation over less organised methods (Boukhelifa et al., 2012). This result is consistent with prior research, where opacity was shown to be the most popular method to visualise inconsistent data from fields outside of educational reporting (Kinkedey, MacEachren and Schiewe, 2014). This could be because students who are non-expert users in interpreting visualisation tended to prefer the intrinsic technique, which does not require the addition of any extra objects to indicate the misalignment in the display, compared with all other visual variables (which were considered extrinsic techniques). Knowing the preferred method to visualise alignment in OLMs could provide guidelines to system designers about the preferred approach that may be beneficial to students.

The third research question of this thesis is 'How is information presentation format important in adding the alignment information in OLMs?' The experimental study in Chapter 7 answers this research question. Researchers argue that the representation method of the learner model (i.e., visual or textual) may influence students' metacognitive skills (Burkett and Azevedo, 2012). Thus, alignments were represented using two methods: graphical and textual. The graphical representation showed the alignment using students' preferred method (opacity), explored in Chapter 4. The graphical representation used the skill meter, which was the most preferred visualisation when the learner model provides different views (Duan, Mitrovic and Churcher, 2010). The second method was the text-based OLM, which has been used widely in prior OLMs as a means to support students'

self-assessment. Providing more detailed information in the learner models, students preferred a fine-grained skill meter (Guerra et al., 2018) and preferred a detailed text-based representation among bar graphs (Lazarinis and Retalis 2007). Taking all of these findings into consideration, the study investigated the impact of the representation method of the alignment on students' bias of confidence judgement on both low- and high-achieving students. The results showed that students using the graphical-based OLM benefited from visualising alignment, and this reflected an increase on their performance especially for low-achieving students. Both high- and low-achieving students who used the graphical-based (expandable model) had more accurate confidence judgement. In contrast, low-achieving students using the text-based OLM needed more support to assess their confidence and their performance, compared to the high-achieving students from the text-based OLM. Students' perceptions in using the system showed that low-achieving students were more motivated to do more activities using the skill meter than the text-based approach, where students felt less motivation to engage with the system. This shows how representing the alignment visually was more beneficial for low-achieving students who need more support. The findings of this study reflected the importance of the representation method when considering designing 'systems that care' (du Boulay et al., 2010) to support students self-regulated learning.

### **8.3 Contribution**

This thesis offers several contributions to the literature of Intelligent Tutoring Systems and Open Learner Models. The first contribution is the OLMlets design and

implementation that I have built to design and visualise model alignment between students' confidence on the correctness of their answers and students' performance, based on the computer evaluation using different representations of alignment to fulfil the aims of this thesis. Students' preferred method to visualise alignment using visual variables was conducted outside of educational reporting. This thesis indicated another contribution; that students preferred the opacity of the colour as their preferred method among other five visual variables, which shows that students preferred the intrinsic technique that provides the overall picture, rather than extrinsic techniques that provides detailed information in the context of OLMs. This result provides guidelines to system designers in educational contexts when students' preferred method is needed.

The third contribution of this thesis is that visualising alignment in a simple OLM enhanced students' confidence gain and motivated students to interact with the system as a result, to align their confidence to their performance.

The fourth contribution of this thesis is that the study suggests that incorporating learners' confidence in the correctness of their answers and students' performance based on the computer evaluation into an OLM particularly benefited the low-achieving students to support their confidence judgement (their confidence relation to their performance).

The fifth contribution showed the importance of how information is presented for Open Learner Model and Intelligent Tutoring System designers, when considering supporting confidence judgement through alignment. Using a graphical

representation of alignment seemed to support students' confidence judgement better than using a text-based Open Learner Model.

#### **8.4 Limitations and Future Work**

While the experimental studies were based on voluntary use by students in a natural class setting (which had a benefit for student learning), this approach limited the number of students who participated in the study. Despite this, the findings were promising and showed that visualising alignment can have an impact on students' confidence and on their behaviour in using the system.

The OLMlets system was available to students at all times; thus, I could not control the effect that students may answer questions with their peers, or look for the answer in their lecture notes. However, system availability allowed students to schedule their own time to use the system and prepare for their class exam. Further, based on the interview conducted, for some students who participated in studies 3 and 4, they were motivated to test their knowledge to prepare for their actual class exam (see Appendix 14, students interview quotes).

One of the limitations of the design of the skill meter visualisation used in OLMlets is that the colours of the skill meter were not designed to work with colour blindness, which could have an adverse effect for students suffering from colour blindness, by misinterpreting the visualisation. Thus, future work should redesign the skill meter visualisation to provide different colour options to allow the student to select different colours, as required.

The study explored in Chapter 7 represents knowledge in the expandable model, which is different from the text-based model. However, both representations use the same modelling process to construct the learner model. The visual display could offer more information based on students' interpretation of the visual display, compared to the text-based approach, which requires more searching within the text (Cox, 1999). This study also provided the graphical OLM using two viewing options (combined and expanded), whereas the text-based OLM only had the expanded view. The expandable model approach that has two methods to visualise alignment (combined and expanded) was taken because some students could have experienced some difficulty in interpreting opacity levels, and research within this thesis showed that students benefited from the two options (combined and expanded) for viewing the model (discussed in Chapter 5).

In this thesis, investigating the impact of visualising alignment was evaluated using a simple form of learner model visualisation (skill meters). Future work could investigate the impact of visualising alignment using more complex views, such as a concept map, radar plot, or tree map. Further, implementing different visual variables used on the survey analysis described in Chapter 4 (in the context of OLMs), and providing students the ability to choose their preferred method to visualise alignment could guide system designers on practical implementation.



# List of References

- Ahmad N., Bull, S. (2008). Do Students Trust their Open Learner Models? In: Nejdil, W., Kay, J., Pu, P., Herder, E. (eds.) International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems, LNCS, vol. 5149, pp. 255--258. Springer, Berlin
- Alotaibi, M. & Bull, S. (2012). Combining Facebook and Open Learner Models to Encourage Collaborative Learning, in G.Biswas, L-H.Wong, T.Hirashima & W.CHEN (eds.) *Proceedings of the 20<sup>th</sup> International conference on Computers in Education*, National Institute of Education, Singapore, pp. 197-201.
- Azevedo, R., & Aleven, V. (eds.), International Handbook of Metacognition and Learning Technologies. *International Handbooks of Education*, 26, Springer.
- Azevedo, R., & Witherspoon, A. M. (2009). Self-regulated learning with hypermedia. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (eds.), *Handbook of metacognition in education* pp. 319-339. Mahwah, NJ: Routledge.
- Azevedo, R., A. Johnson, A. Chauncey, A. Graesser, B. Zimmerman, and D. Schunk. (2011). Use of hypermedia to convey and assess self-regulated learning. *Handbook of self-regulation of learning and performance*, pp.102–121.
- Azevedo, R., Witherspoon, A., Chauncey, A., Burkett, C., & Fike, A. (2009). MetaTutor: A MetaCognitive tool for enhancing self-regulated learning. In R.

- Pirrone, R. Azevedo, & G. Biswas (eds.), *Proceedings of the AAAI Fall Symposium on Cognitive and Metacognitive Educational Systems*, pp. 14–19.
- Bertin, J. (1967/1983). *Semiology of graphics: diagrams, networks, maps* (translated by William J. Berg). University of Wisconsin Press.
- Bloom, B. S., Englehart, M. D., Furst, E. J., Hill, W.H., & Krathwohl, D. R., (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook 1: Cognitive domain*. New York: McKay.
- Bonneau, G. P., Hege, H. C., Johnson, C. R., Oliveira, M. M., Potter, K., Rheingans, P., & Schultz, T. (2014). Overview and state-of-the-art of uncertainty visualization. In *Scientific Visualization*, Springer, London, pp.3-27.
- Bodily, R., Kay, J., Alevan, V., Davis, D., Jivet, I., Xhakaj, F., & Verbert, K. (2018). Open Learner models and learning analytics dashboards: a systematic review. In *Proceedings of the Eight International Conference on Learning Analytics & knowledge*. ACM, Sydney, Australia, pp.41-50.
- Boud, D. (1989). The Role of Self-Assessment in Student Grading, Assessment & Evaluation in *Higher Education*, 14(1), pp.20–30
- Boukhelifa, N., Bezerianos, A., Isenberg, T., & Fekete, J. D. (2012). Evaluating sketchiness as a visual variable for the depiction of qualitative uncertainty. *IEEE Transactions on Visualisation and Computer Graphics*. 18(12), pp. 2769–2778.

- Brusilovsky P., Somyürek S., Guerra J., Hosseini R., & Zadorozhny V. (2015). The value of social: Comparing open student modelling and open social student modelling. *International conference on user modelling, adaptation, and personalization*. Dublin, Ireland, pp.44–55.
- Brusilovsky, P., Hsiao, I.-H. & Folajimi, Y. (2011). QuizMap: Open Social Student Modelling and Adaptive Navigation Support with Tree Maps, *Towards Ubiquitous Learning* SE - 7, (6964), pp.71–82.
- Bull, S. & Al-Shanfari, L. (2015). Negotiating Individual Learner Models in Contexts of Peer Assessment and Group Learning, in J.Boticario & K. Muldner (eds.) in *ISLG2015 fourth Workshop on Intelligent Support for Learning in Groups, Proceedings of the Workshop at the 17<sup>th</sup> International Conference on Artificial Intelligence in Education, AIED 2015, CEUR, Madrid, Spain*, pp. 1-6.
- Bull, S. & Mabbott, A. (2006). 20000 Inspections of a Domain-Independent Open Learner Model with Individual and Comparison Views, in Ikeda, M., Ashley, K. D., and Chan, T.-W. (eds.) *Intelligent Tutoring Systems*. Jhongli, Taiwan: Springer Berlin Heidelberg, pp.422–432.
- Bull, S., Johnson, M. D., Masci, D., & Biel, C. (2016a). Integrating and Visualising diagnostic Information for the Benefit of Learning, P. Reimann, S. Bull, M.Kickmeier-Rust, R.K. Vatrapsu & B. Wasson (eds.) *Measuring and Visualizing Learning in the Information-rich classroom*. Routledge/Taylor and Francis, pp.167–180.

- Bull, S., Ginon, B., Boscolo, C. & Johnson, M. (2016b). Introduction of Learning Visualisations and Metacognitive Support in a Persuadable Open Learner Model. Proceedings of LAK16 6th International Conference on Analytics and Knowledge, pp. 30–39.
- Bull, S. & Gardner, P. (2010). Raising learner awareness of progress towards UK-SPEC learning outcomes. *Engineering Education*, 5(1), pp. 11–22.
- Bull, S., Jackson, T. & Lancaster, M. (2010). Students' interest in their misconceptions in first-year electrical circuits and mathematics courses. *International Journal of Electrical Engineering Education*, 47(3), pp. 307–318.
- Bull, S., Johnson, M. D., Demmans Epp, C., Masci, D., Alotaibi, M., & Girard, S. (2014). Formative assessment and meaningful learning analytics. *International Conference in Advanced Learning Technologies (ICALT), IEEE*, Athens, Greece, pp.327-329.
- Bull, S. & Kay, J. (2013). Open Learner Models as Drivers for Metacognitive Processes, In R. Azevedo and V. Aleven (eds.), *International Handbook of Metacognition and Learning Technologies*. Springer-Verlag New York, pp. 349-365.
- Bull, S. & Kay, J. (2016). SMILI ☺ : a Framework for Interfaces to Learning Data in Open Learner Models, Learning Analytics and Related Fields. *International Journal of Artificial Intelligence in Education*, 26(1), pp. 293–331.

- Bull, S. & McKay, M. (2004). An Open Learner Model for Children and Teachers: Inspecting Knowledge Level of Individuals and Peers. *Intelligent Tutoring Systems (ITS'2004)*, pp. 646–655.
- Bull, S. & Pain, H. (1995). “Did I Say what I Think I Said, and Do You Agree with Me?”: Inspecting and Questioning the Student Model, in Greer, J. (ed.) *Proceedings of World Conference on Artificial Intelligence and Education*. Charlottesville VA, USA: Association for the Advancement of Computing in Education, pp. 501–508.
- Bull, S., Pain, H. & Brna, P. (1995). Mr.Collins: A collaboratively constructed, inspectable student model for intelligent computer assisted language learning, *Instructional Science*, 23(1), pp.65-87.
- Butler, D. L. & Winne, P. H. (1995). Feedback and self-regulated learning: A Theoretical Synthesis, *Review of Educational Research*, 65(3), pp. 245–281.
- Burkett, C., & Azevedo, R. (2012). The effect of multimedia discrepancies on metacognitive Judgements. *Computers in Human Behavior*, 28(4), pp.1276-1285.
- Carr, B., and Goldstein, I. (1977). *Overlays: A theory of modelling for computer aided instruction* (AI Memo 406). Cambridge, MA: Massachusetts Institute of Technology, AI Laboratory.
- Callender, A. A., Franco-Watkins, A. M. & Roberts, A. S. (2016). Improving metacognition in the classroom through instruction, training, and feedback, *Metacognition and Learning*, 11(2), pp. 215–235.

- Cohen, J., (1988). *Statistical power analysis for the behavioural sciences*. Hillsdale, NJ: Erlbaum
- Conejo, R., Trella, M., Cruces, I. and Garcia, R. (2011). INGRID: A web service tool for hierarchical open learner model visualization, In: L. Ardissono and T. Kuflik (eds.) *Advances in User Modelling. UMAP 2011. in Lecture Notes in Computer Science (7138)*, Springer, Berlin, Heidelberg, pp. 406-409.
- Corbett, A.T., Anderson, J.R., (1995). Knowledge tracing: Modeling the acquisition of procedural knowledge, *User Modelling and User-Adapted Interaction* 4, pp. 253-278.
- Corbett, A. T. & Bhatnagar, A. (1997). Student modelling in the ACT programming tutor: Adjusting a procedural learning model with declarative knowledge, in Jameson, A., Paris, C., and Tasso, C. (eds.) *User Modelling*. Sardinia, Italy: Springer Vienna, pp. 243–254.
- Correa, C.D., Chan, Y.H., and Ma, K.L. (2009). A framework for uncertainty-aware visual analytics. *VAST 09 - IEEE Symposium on Visual Analytics Science and Technology*, Proceedings, pp. 51-58.
- Cox, R. (1999). Representation construction, externalised cognition and individual differences, *Learning and instruction*, 9, pp. 343–363.
- Creswell, J. W., & Plano Clark, V. L. (2007). *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage.

- Demmans Epp, C. & Bull, S. (2015). Uncertainty Representation in Visualizations of Learning Analytics for Learners: *Current Approaches and Opportunities*, *IEEE Transactions on Learning Technologies*, 1382(3), pp. 242–260.
- Dewey, J. (1938). *Experience and education*. New York: Simon and Schuster.
- DiFrancesca, D., Nietfeld, J. L. & Cao, L. (2016). A comparison of high and low achieving students on self-regulated learning variables, *Learning and Individual Differences*. Elsevier Inc., (45), pp. 228–236.
- Dimitrova, V. (2003). STyLE-OLM: Interactive Open Learner Modelling. *International Journal of Artificial Intelligence in Education (IJAIED)*, (13), pp. 35–78.
- Duan, D., Mitrovic, A. & Churcher, N. (2010). Evaluating the Effectiveness of Multiple Open Student Models in EER-Tutor, In: WONG, S. L. (ed.) *Proceedings of the 18th International Conference on Computers in Education*. Putrajaya, Malaysia: Asia-Pacific Society for Computers in Education, pp. 86–88.
- du Boulay, B., Avramides, K., Luckin, R., Martinez-Miron, G. R., & Carr, A. (2010). Towards systems that care: A conceptual framework based on motivation, metacognition, and affect. *International Journal of Artificial Intelligence in Education*, (20), pp. 197-229.
- Duffy, M. C., & Azevedo, R. (2015). Motivation matters: Interactions between achievement goals and agent scaffolding for self-regulated learning within an intelligent tutoring system. *Computers in Human Behavior*, (52), pp. 338-348.

- Dunlosky, J., & A. Rawson, K. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction*, 22 (4), pp. 271-280.
- Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., & Kruger, J. (2008). Why the unskilled are unaware: further explorations of (absent) self-insight among the incompetent. *Organizational Behavior and Human Decision Processes*, 105, pp.98–121.
- Feyzi-Behnagh, R., Azevedo, R., Legowski, E., Reitmeyer, K., Tseytlin, E., & Crowley, R. S. (2014). Metacognitive scaffolds improve self-Judgements of accuracy in a medical intelligent tutoring system. *Instructional science*, 42(2), pp. 159-181.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry., *American Psychologist*, 34(10), pp. 906–911.
- Gama, C. (2004). Metacognition in Interactive Learning Environments: The Reflection Assistant Model. In J. C. Lester, R. m. Vicari, & F. Paraguacu (eds.), 7<sup>th</sup> International Conference on Intelligent Tutoring Systems, ITS 2004, (Vol. Lecture Notes in Computer Science 3220, pp. 668-677.
- Gerharz, L. & Pebesma, E. (2009). Usability of interactive and non-interactive visualisation of uncertain geospatial information, *Geoinformatik 2009 Konferenzband*, (MacEachren), pp. 223–230.



- Greene, J. A., & Azevedo, R. (2007). A Theoretical Review of Winne and Hadwin's Model of Self-Regulated Learning: New perspectives and Directions. *Review of Educational Research*, 77 (3), pp. 334-372.
- Griethe, H. & Schumann, H. (2006). The visualization of uncertain data: Methods and problems, *Proceedings of SimVis 2006*, SCS Publisher House, pp. 143–156.
- Guerra, J., Schunn, C. D., Bull, S. & Barria-Pineda, J. (2018). Navigation support in complex open learner models: assessing visual design alternatives Julio, *New Review of Hypermedia and Multimedia*. Taylor & Francis, 0(0), pp. 1–29.
- Hacker, D. J., Bol, L., Horgan, D. D., & Rakow, E. A. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology*, 92(1), pp.160–170.
- Hartley, D. & Mitrovic, A. (2002). Supporting learning by opening the student model, in Cerri, S. A., Gouardères, G., & Paraguaçu, F. (eds.), *Intelligent Tutoring Systems*. Biarritz, France and San Sebastian, Spain: Springer Berlin Heidelberg, pp. 453–462.
- Holt, P., Dubs, S., Jones, M., Greer, J. (1994). The state of student modelling. In Greer, J., McCalla, G. (eds.) *Student Modelling: The Key to Individualized Knowledge-Based Instruction*, Springer, New York, pp. 3–39.

- Huff, J. D., & Nietfeld, J. L. (2009). Using strategy instruction and confidence Judgements to improve metacognitive monitoring. *Metacognition and Learning*, 4(2), pp.161–176.
- Hull, A., & du Boulay, B. (2009). Scaffolding Motivation and Metacognition in Learning Programming. In V. Dimitrova, R. Mizoguchi, B. du Boulay, & A. Grasser (eds.), *Artificial Intelligence in Education. Building Learning Systems that Care: from knowledge Representation to Affective Modelling* (Vol. Frontiers in Artificial Intelligence and Applications 200, pp. 755-756.
- Jackson, S. a. & Kleitman, S. (2014). Individual differences in decision-making and confidence: Capturing decision tendencies in a fictitious medical test', *Metacognition and Learning*, 9(1), pp. 25–49.
- Kay, J. (1997). Learner Know Thyself: Student Models to Give Learner Control and Responsibility, *International Conference on Computers in Education*, (10), pp. 17–24.
- Kay, J., Li, L. & Fekete, A. (2007). Learner Reflection in Student Self-assessment, in Mann, S. & Simon (eds.) *Proc. Ninth Australasian Computing Education Conf.* Darlinghurst, Australia: Australian Computer Society, Inc., pp. 89–95.
- Kelemen, W. L. (2000). Metamemory Cues and Monitoring Accuracy: Judging What You Know and What You Will Know, *Journal of Educational Psychology*, 92(4), pp. 800–810.
- Kerly, A., & Bull, S. (2008). Children's interactions with inspectable and negotiated learner models. In B. P. Woolf, E. Aimeur, R. Nkambou, & S. Lajoie

- (Eds.), *Intelligent Tutoring Systems: 9th International Conference*. Berlin, Heidelberg: Springer. pp. 132–14.
- Kerly, A., Ellis, R. & Bull, S. (2008). CALMsystem: A Conversational Agent for Learner Modelling, *Knowledge-Based Systems*, 21 (3), pp. 238–246.
- Kerly, A., Hall, P. & Bull, S. (2007). Bringing chatbots into education: Towards natural language negotiation of open learner models, *Knowledge-Based Systems*, (20), pp. 177–185.
- Kinkeldey, C., MacEachren, A. M. & Schiewe, J. (2014). How to assess visual communication of uncertainty? A systematic review of geospatial uncertainty visualisation user studies, *Cartographic Journal*, 51(4), pp. 372–386.
- Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to Judgements of learning, *Journal of Experimental Psychology: General*, 126(4), pp. 349–370.
- Kulhavy, R. W. & Stock, W. A. (1989). Feedback in Written Instruction: The Place of Response Certitude, *Educational Psychology Review*, 1(4), pp. 279–308.
- Lan, C.H., Tseng, C.C. & Lai, K.R., 2008, July. Developing a negotiation-based intelligent tutoring system to support problem solving: A case study in role-play learning. In 2008 Eighth IEEE International Conference on Advanced Learning Technologies, pp. 356-360.
- Law, C. Y., Grundy, J., Cain, A., & Vasa, R. (2015). A preliminary study of open learner model representation formats to support formative assessment, in

*Proceedings of the IEEE 39th Annual International Computers, Software & Applications Conference (COMPSAC)*, pp. 887– 892.

Lazarinis, F. and Retalis, S., (2007). Analyze me: Open learner model in an adaptive web testing system, *International Journal of Artificial Intelligence in Education*, 17, pp. 255–271.

Long, Y. & Alevan, V. (2013a). Active Learners: Redesigning an Intelligent Tutoring System to Support Self-Regulated Learning, in Hernandez-Leo, D. (ed.) *Proceedings of the 8th European Conference on Technology Enhanced Learning*, EC-TEL 2013. New York: Springer, pp. 490–495.

Long, Y., & Alevan, V. (2013b). Supporting students' self-regulated learning with an open learner model in a linear equation tutor. In H. C. Lane, K. Yacef, J. Mostow, & P. Pavlik (eds.), *Artificial intelligence in education*, Berlin: Springer, pp. 219–228.

Long, Y. & Alevan, V. (2013c). Skill Diaries: Improve Student Learning in an Intelligent Tutoring System with Periodic Self-Assessment, in Lane, H. C. et al. (eds.), *Artificial Intelligence in Education*. Memphis, TN, USA: Springer Berlin Heidelberg, pp. 249–258.

Mabbott, A. & Bull, S. (2004). Alternative views on knowledge: Presentation of open learner models, in Lester, J. C., Vicari, R. M., & Paraguaçu, F. (eds), *Intelligent Tutoring Systems*. Maceió, Alagoas, Brazil: Springer Berlin Heidelberg, pp. 689–698.

- Mabbott, A. & Bull, S. (2006). Student preferences for editing, persuading, and negotiating the open learner model, in *Lecture Notes in Computer Science* (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 481–490.
- MacEachren, A. M. (1992). Visualizing uncertain information, *Cartographic Perspective*, 13(13), pp. 10–19.
- Maki, R. H., Shields, M., Wheeler, A. E., & Zacchilli, T. L. (2005). Individual Differences in Absolute and Relative Metacomprehension Accuracy. *Journal of Educational Psychology*, 97(4), pp.723-731.
- Mathan, S. A. & Koedinger, K. R. (2005). Fostering the Intelligent novice: Learning from errors with metacognitive tutoring, *Educational Psychologist*, 40(4), pp. 257–265.
- Miller, T. M. & Geraci, L. (2011). Training metacognition in the classroom: The influence of incentives and feedback on exam predictions, *Metacognition and Learning*, 6(3), pp. 303–314.
- Mitrovic, A. & Martin, B. (2007). Evaluating the Effect of Open Student Models on Self-Assessment, *International Journal of Artificial Intelligence in Education*, 17(2), pp. 121–144.
- Morrison, J.L. A theoretical framework for cartographic generalization with the emphasis on the process of symbolization. *International Yearbook of Cartography*, 14, 115-127, 1974. Cited in MacEachren, A.M., Roth, R.E., O'Brien, J., Swingley, D. and Gahegan, M. (2012). *Visual semiotics and*

*uncertainty visualization: an empirical study*. IEEE Trans., Visualization and Computer Graphics, 18 (12), pp. 2496–2505

Nietfeld, J. L., Cao, L. & Osborne, J. W. (2005). Metacognitive Monitoring Accuracy and Student Performance in the Postsecondary Classroom, *Journal of Experimental Education*, 74(1), pp. 7-28.

Nietfeld, J., & Schraw, G. (2002). The effect of knowledge and strategy training on monitoring accuracy. *The Journal of Educational Research*, 95, pp. 131–142.

Nwana, H. (1990). Intelligent tutoring systems: an overview, *Artificial Intelligence Review*, 4(4), pp. 251–277.

Pintrich, P. R., Wolters, C. A & Baxter, G. P. (2000). Assessing Metacognition and Self-Regulated Learning, *Issues in the Measurement of Metacognition*, pp. 43–97.

Renner, C. H. & Renner, M. J. (2001). But I Thought I Knew That: Using Confidence Estimation as a Debiasing Technique to Improve Classroom Performance, *Applied Cognitive Psychology*, 15(1), pp. 23–32.

Roll, I., Alevan, V., McLaren, B. M., & Koedinger, K. R. (2011). Metacognitive practice makes perfect: Improving students' self-assessment skills with an intelligent tutoring system. In *International Conference on Artificial Intelligence in Education*, Springer, Berlin, Heidelberg, pp. 288-295.

- Royer, J. M., Cisero, C. A. & Carlo, M. S. (1993). Techniques and Procedures for Assessing Cognitive Skills, *Review of Educational Research*, 63(2), pp. 201–243.
- Schon, D., (1983). *The Reflective Practitioner*. Basic Books, USA.
- Schraw, G. (2009). A conceptual analysis of five measures of metacognitive monitoring, *Metacognition and Learning*, 4(1), pp. 33–45.
- Schraw, G. & Dennison, R. S. (1994). Assessing Metacognitive Awareness, *Contemporary Educational Psychology*, pp. 460–475.
- Self, J. (1990). Bypassing the intractable problem of student modelling, *Intelligent Tutoring Systems: At the crossroads of artificial intelligence and education*, (41), pp. 107–123.
- Shapiro, S., & Wilk, M. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, 52(3/4), pp.591-611.
- Slocum, T. A., Cliburn, D. C., Feddema, J. J., & Miller, J. R. (2003). Evaluating the usability of a tool for visualizing the uncertainty of the future global water balance. *Cartography and Geographic Information Science*, 30(4), pp.299-317.
- Stampfer, E., Long, Y., Aleven, V., & Koedinger, K. R. (2011). Eliciting intelligent novice behaviors with grounded feedback in a fraction addition tutor. In *International Conference on Artificial Intelligence in Education*, Springer, Berlin, Heidelberg. pp. 560-562.

- Stankov, L., Kleitman, S., & Jackson, S. A. (2014). Measures of the trait of confidence. In G. J. Boyle, D. H. Saklofske, & G. Matthews (eds.), *Measures of personality and social psychological constructs*, Academic Press, London, UK, pp. 158–189.
- Suleman, R.M., Mizoguchi, R. & Ikeda, M. (2015). Negotiation-Driven Learning, in C. Conati, N. Heffernan, A. Mitrovic & M.F. Verdejo (eds.), *Artificial Intelligence in Education*, Springer International Publishing, pp. 470-479
- Tatsuoka, K.: Rule space: An approach for dealing with misconceptions based on item response theory. *Journal of Educational Measurement*. 20(4) (1983) 345-354.
- Thiede, K. W., Anderson, M. C. M. & Therriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts, *Journal of Educational Psychology*, 95(1), pp. 66–73.
- Thomson, J., Hetzler, E., MacEachren, A., Gahegan, M., & Pavel, M. (2005, March). A typology for visualizing uncertainty. In *Visualization and Data Analysis*, *International Society for Optics and Photonics*. (5669), pp. 146-158).
- Thurstone, L. (1927). A Law of Comparative Judgement, *Psychological Review*, 101(2), pp. 266–270.
- Tobias, S. & Everson, H. T. (2002). Knowing what you know and what you don't: *Further research on metacognitive knowledge monitoring*, *College Board Report No. 2002-3*. New York.



- Tobias, S., & Everson, H. T. (2009). The importance of knowing what you know: A knowledge monitoring framework for studying metacognition in education. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (eds.), *Handbook of metacognition in education*, New York, NY: Routledge, pp. 107-128.
- Trevors, G., Duffy, M., & Azevedo, R. (2014). Note-taking within MetaTutor: Interactions between an intelligent tutoring system and prior knowledge on note-taking. *Educational Technology Research & Development*, (62), pp. 507-528
- Uther, J. B. (2001). On the Visualisation of Large User Models in Web Based Systems, PhD, University of Sydney, Sydney, Australia.
- Vanlehn, K. (2006). The behavior of tutoring systems. *International Journal of Artificial Intelligent in Education*. 16(3). pp. 227–265.
- VanZile-Tamsen, C. & Livingston, J., (1999). The differential impact of motivation on the self-regulated strategy use of high- and low-achieving college students. *Journal of College Student Development*, 40(1), pp.54–60
- Ware, C. (2004). *Information visualization: perception for design*. San Francisco: Morgan Kaufmann.
- Woolf, B. (2009). *Building intelligent interactive tutors: Student-centered strategies for revolutionizing e-learning*. Morgan Kaufmann, Burlington, pp. 9-94.

Zapata-Rivera, J. D. & Greer, J. E. (2004). Interacting with Inspectable Bayesian Student Models, *International Journal of Artificial Intelligence in Education*, (14), pp. 127–163.

Zimmerman, B. J. (1990). Self-Regulated Learning and Academic Achievement: An Overview, *Educational Psychologist*, 25(1), pp. 3–17.

Zuk, T. & Carpendale, S. (2006). Theoretical analysis of uncertainty visualizations, *Proc. SPIE*, (6060), pp. 606007–606014.

# Appendix 1: Paired Comparison Survey (Chapter 4)

STUDENT ID \_\_\_\_\_

This questionnaire is designed to get feedback about your preference in the design of the visualisation of the system model. The system model is constructed based on your responses to multiple choice questions. After each question students are asked to rate their confidence about the correctness of the answer. The visualisation of the system model combines the data of the system's beliefs about your answer and the student's confidence about the correctness of the answer showing different levels of agreement between the system and the student.

Please specify your Gender:  Female  Male

Have you ever used Teaching Adaptive System?  Yes  No

If Yes, the name of the Teaching Adaptive System is  OLMlets OLM  Others

Which visualisation you prefer between Figure 1 and Figure 2?

**Please interpret the legend of the figures to understand the levels of agreement represented in the skill meters. Please answer the following questions by ticking the box beside the figure. The data will be stored anonymously.**

Figure 1		Figure 2	
<p><b>Levels of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□	<p><b>Levels of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□
<p><b>Levels of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System Model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□	<p><b>Level of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System Model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□
<p><b>Levels of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System Model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□	<p><b>Levels of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System Model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□
<p><b>Levels of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System Model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□	<p><b>Levels of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□
<p><b>Level of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System Model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□	<p><b>Levels of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□

<p><b>Levels of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□	<p><b>Levels of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System Model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□
<p><b>Levels of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System Model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□	<p><b>Levels of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System Model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□
<p><b>Level of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System Model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□	<p><b>Levels of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□
<p><b>Levels of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□	<p><b>Levels of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System Model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□
<p><b>Level of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System Model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□	<p><b>Levels of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System Model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□

<p><b>Levels of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□	<p><b>Levels of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System Model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□
<p><b>Level of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System Model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□	<p><b>Levels of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System Model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□
<p><b>Levels of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□	<p><b>Levels of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System Model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□
<p><b>Levels of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System Model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□	<p><b>Levels of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System Model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□
<p><b>Levels of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□	<p><b>Levels of agreement between your responses and your confidence</b></p> <p>known misconception unknown</p> <p>Full agreement   </p> <p>High agreement   </p> <p>Medium agreement   </p> <p>Low agreement   </p> <p><b>System Model</b> <b>Topic</b></p> <p> The Algebra of Complex Numbers</p> <p> Point Representation of Complex Numbers</p> <p> Vectors and Polar Forms</p>	□

## Appendix 2: Paired Comparison Survey: Student data (SPSS file)

	Gender	Used OLM	Arrangement	Opacity	Size	Orientation	Numerosity	Added marks
A1	m	1	1	5	3	4	2	0
A2	m	1	0	4	1	4	4	2
A3	m	1	0	2	3	2	3	5
A4	m	1	0	5	3	2	2	3
A5	f	1	3	5	2	1	4	0
A6	f	1	0	4	1	4	4	2
A7	f	1	1	5	0	2	3	4
A8	m	1	1	5	0	2	4	3
A9	m	1	2	3	4	1	5	0
A10			0	1	4	3	5	2
A11	f	1	1	2	4	4	1	3
A12	m	1	1	5	0	3	2	4
A13	m	1	0	4	3	4	3	1
A14	m	1	2	4	2	3	0	4
A15	f	1	3	1	2	3	2	4
A16	m	1	1	2	4	5	1	2
A17			0	2	3	1	4	5
A18	m	1	2	0	4	5	1	3
A19	m	1	1	5	3	2	4	0
A20	m	1	1	5	3	2	0	4
A21			2	4	3	1	5	0
A22	m	1	0	5	3	3	3	1
A23	m	1	0	5	4	1	3	2
A24	m	1	0	3	2	4	1	5
A25	m	1	0	3	1	3	5	3
A26	m	1	1	5	0	2	3	4
A27	m	1	2	5	4	3	1	0
A28	m	1	2	3	2	5	1	2
A29	m	1	0	5	3	2	1	4
A30	f	1	0	5	1	2	3	4
A31	f	1	0	2	2	2	4	5

A32	m	1	4	5	2	1	3	0
A33	f	1	0	5	4	3	1	2
A34	f	1	1	5	1	4	3	1
A35			0	3	5	3	3	1
A36	m	1	1	5	3	4	1	1
A37	m	1	1	2	2	1	5	4
A38	m	1	0	5	4	2	3	1
A39	m	1	0	2	1	3	4	5
A40	f	1	0	4	1	2	3	5
A41	f	1	0	5	1	2	4	3
A42	f	0	2	0	2	4	4	3
A43	f	1	0	4	5	1	2	3
A44	f	1	0	5	2	1	4	3
A45	m	0	0	5	2	2	3	3
A46	m	1	0	5	1	3	3	3
A47	m	0	2	3	4	3	1	2
A48	m	0	0	5	1	2	4	3
A49	m	0	0	5	4	1	2	3
A50	m	0	1	3	3	4	3	1
A51	m	0	2	4	2	4	1	2
A52	m	0	0	5	3	4	1	2
A53	f	0	0	4	2	2	5	2
A54	m	0	0	5	2	4	3	1
A55	m	0	0	5	3	4	2	1
A56	m	0	1	5	4	0	2	3
A57	m	0	3	3	3	2	1	3
A58	m	0	4	2	1	5	2	1
A59	m	0	2	3	2	4	1	3
A60	m	0	0	5	2	3	1	4
A61	f	1	0	3	2	1	5	4
A62	f	0	2	3	4	3	1	2
A63	m	0	0	5	1	2	4	3
A64	f	0	1	5	1	1	4	3
A65	m	0	1	3	2	1	3	5
A66	m	0	0	5	1	4	3	2
A67	m	0	0	5	4	2	2	2



# Appendix 3: Pre-questionnaire (Chapter 5)

STUDENT ID \_\_\_\_\_

This questionnaire is designed to get feedback about your learning skills. Please answer the following questions based on your experiences.

The data will be stored **anonymously**.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
<b>Your Learning</b>					
Doing homework helps me see how much I know.					
Taking tests helps me to identify gaps in my knowledge.					
Taking tests helps me to identify my misconceptions.					
Doing homework helps me to monitor my learning.					
I use tests and homework to plan for my learning.					
I let the teacher or my parents plan for my learning.					
I plan for my own learning.					
<b>Your confidence</b>					
I am good at assessing my confidence level.					
I try to increase my knowledge when my confidence is low.					
I try to increase my knowledge when my confidence is high.					
My confidence increases when I do well on a test or homework assignment.					
I feel my confidence decreases when I do poorly on a test.					

I know how well I will do before I write a test.					
--	--	--	--	--	--

**Comments:**

Is there anything else that you would like to tell us:

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

# Appendix 4: Pre-questionnaire students' responses (Chapter 5)

(SPSS file): all groups

user_id	OLMlets use	Pre-Groups	Q1 Learning	Q2 Learning	Q3 Learning	Q4 Learning	Q5 Learning	Q6 Learning	Q7 Learning	Q1 Confidence	Q2 Confidence	Q3 Confidence	Q4 Confidence	Q5 Confidence	Q6 Confidence
103238	0	1	5	4	4	4	3	2	4	4	5	5	5	2	2
111414	0	1	5	5	5	5	5	1	5	3	5	4	5	3	5
106022	0	1	4	3	4	4	3	2	4	3	4	3	4	4	2
107435	0	1	4	5	4	3	4	2	5	4	3	4	5	4	2
113333	0	1	4	4	4	4	4	3	4	4	5	5	5	5	4
106808	0	1	3	4	4	5	5	5	5	5	5	5	5	5	5
108357	0	1	4	4	3	4	3	2	4	1	4	4	3	3	2
100868	0	1	4	4	4	4	3	2	5	4	4	5	4	4	3
102321	0	1	5	5	5	4	4	2	4	4	4	2	4	2	4
101024	0	1	5	4	4	4	4	3	5	5	4	4	5	4	4

102166	0	2	3	2	2	4	4	2	5	4	5	3	5	5	4
94265	0	2	5	4	3	4	2	1	3	3	1	2	3	4	4
100788	0	2	5	4	5	3	4	2	4	4	4	3	5	3	3
103271	0	2	5	4	3	4	5	4	5	4	5	3	5	4	2
103236	0	2	5	4	3	5	4	2	5	3	5	5	3	5	2
94022	0	2	4	3	4	5	3	2	3	3	4	3	5	5	2
106814	0	2	3	4	4	3	4	4	4	3	4	4	3	3	4
100800	0	2	4	3	4	3	4	3	4	4		4	4	4	3
101203	0	2	4	4	3	5	5	4	3	4	4	3	3	4	5
100711	0	3	5	4	4	4	3	1	5	4	5	5	5	4	3
102234	0	3	5	5	5	5	5	4	5	4	5	5	4	3	4
110928	0	3	5	4	3	4	5	2	3	3	4	4	5	5	3
110660	0	3	5	4	4	3	5	3	5	5	4	4	5	3	4
113391	0	3	5	4	4	3	4	2	5	5	5	4	5	5	3
107562	0	3	2	5	5	2	3	4	5	3	4	4	4	3	3
106359	0	3	4	4	3	4	5	3	4	4	4	5	5	3	4
107952	0	3	5	4	4	4	4	5	5	3	5	5	5	4	4
101572	1	1	5	4	5	4	3	3	5	3	4	3	5	5	3
100598	1	1	4	5	4	4	5	5	4	4	5	4	5	5	4
103426	1	1	4	4	3	4	4	3	4	4	3	5	5	5	2
101967	1	1	4	3	4	4	5	1	5	3	4	3	3	4	1
107951	1	1	5	5	5	5	4	2	5	4	2	5	5	3	2
107446	1	1	4	4	4	3	3	1	5	2	4	5	5	2	1
100750	1	1	5	4		5	4	4	4	4	4		5	4	4
93002	1	1	5	4		3	3	3	5	4	4	3	3	4	4

101656	1	2	4	5	4	4	3	2	5	4	4	5	5	3	5
96028	1	2	4	4	4	4	4	3	4	4	4	4	4	3	3
96773	1	2	4	3	4	4	3	4	2	5	4	4	5	3	3
101270	1	2	5	4	4	5	4	4	5	4	4	4	4	4	4
103447	1	3	5	5	4	5	5	5	5	4	5	5	5	3	2
102994	1	3	5	4	5	5		2	5	4	4	5	5	4	4
108300	1	3	5	5	4	4	4	2	3	4	5	4	5	5	3
100819	1	3	5	5	5	4	3	2	4	4	5	4	4	3	3
99146	1	3	5	5	5	4	5	2	4	5	5	5	5	5	4
105932	1	3	5	5	5	5	5	5	3	5	5	5	5	5	5
97001	1	3	4	4	5	4	4	4	3	4	4	4	4	4	4

# Appendix 5: Post-questionnaire control group (Chapter 5)

STUDENT ID \_\_\_\_\_

This questionnaire is designed to get feedback on OLMlets. Please answer the following questions based on your experience while using the system. The data will be stored anonymously.

## OLMlets OLM Questionnaire:

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
<b>Your Learning</b>					
OLMlets helped me see how much I know.					
OLMlets helped me to identify gaps in my knowledge.					
OLMlets helped me to identify my misconceptions.					
OLMlets helped me to monitor my learning.					
OLMlets helped me to plan for my learning.					
OLMlets encouraged me to answer more questions.					
<b>Your confidence</b>					
I am good at assessing my confidence level.					
I try to increase my knowledge when the system shows that I had low level of knowledge and I feel that my confidence is low.					
I try to increase my knowledge when the system shows that I had low level of knowledge and I feel that my confidence is high.					
I felt my confidence increase when the system shows that I had a high level of knowledge.					

I felt my confidence decrease when the system shows that I had low level of knowledge.					
Feeling that my confidence level matched my knowledge level encouraged me solve more questions.					
Feeling that my confidence level was different from my knowledge level encouraged me solve fewer questions.					
<b>Comments:</b> Is there anything else (good or bad) that you would like to tell us: ----- ----- ----- ----- ----- ----- ----- ----- ----- ----- ----- ----- ----- ----- ----- -----					

# Appendix 6: Post-questionnaire combined group (Chapter 5)

STUDENT ID \_\_\_\_\_

This questionnaire is designed to get feedback on OLMlets. Please answer the following questions based on your experience while using the system. The data will be stored anonymously.

## OLMlets OLM Questionnaire:

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
<b>Your Learning</b>					
OLMlets helped me to see how much I know .					
OLMlets helped me to identify my gaps in my knowledge.					
OLMlets helped me to identify my misconceptions.					
OLMlets helped me to plan for my learning.					
OLMlets helped me to monitor my learning.					
OLMlets encouraged me to answer more questions.					
<b>Your Confidence</b>					
I am good at assessing my confidence level.					
The system helped me to identify topics where I am under-confident.					
The system helped me to identify topics where I am over-confident.					



I try to increase my knowledge when the system shows that I had low level of knowledge and I feel that my confidence is low.					
I try to increase my knowledge when the system shows that I had low level of knowledge and I feel that my confidence is high.					
I felt my confidence increase when the system shows that I had a high level of knowledge.					
I felt my confidence decrease when the system shows that I had low level of knowledge.					
Feeling that my confidence level matched my knowledge level encouraged me solve more questions.					
Feeling that my confidence level was different from my knowledge level encouraged me solve fewer questions.					
<b>OLMlets Charts</b>					
I can easily notice from the chart, when my confidence and my level of knowledge match completely.					
I can easily notice from the chart, when my confidence and my level of knowledge mostly match.					
I can easily notice from the chart, when my confidence and my level of knowledge match somewhat					
I can easily notice from the chart, when my confidence and my level of knowledge are completely different					
Being shown that there is a difference between my knowledge and confidence is helpful					
Seeing my model let me be more accurate in assessing my confidence level					
Being shown that there is a difference between my knowledge and confidence led me to solve more questions.					
Being shown that there is a difference between my knowledge and confidence led me to answer fewer questions.					
<b>Comments:</b> Is there anything else (good or bad) that you would like to tell us: ----- ----- -----					

# Appendix 7: Post-questionnaire expandable group (Chapter 5)

STUDENT ID \_\_\_\_\_

This questionnaire is designed to get feedback on OLMlets. Please answer the following questions based on your experience while using the system. The data will be stored anonymously.

## OLMlets OLM Questionnaire:

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
<b>Your Learning</b>					
OLMlets helped me to see how much I know .					
OLMlets helped me to identify my gaps in my knowledge.					
OLMlets helped me to identify my misconceptions.					
OLMlets helped me to monitor my learning.					
OLMlets helped me to plan for my learning.					
OLMlets encouraged me to answer more questions.					
<b>Your confidence</b>					
I am good at assessing my confidence level.					
The system helped me to identify topics where I am under-confident.					
The system helped me to identify topics where I am over-confident.					
I try to increase my knowledge when the system shows that I had low level of knowledge and I feel that my confidence is low.					
I try to increase my knowledge when the system shows that I had low level of knowledge and I feel that my confidence is high.					

I felt my confidence increase when the system shows that I had a high level of knowledge.					
I felt my confidence decrease when the system shows that I had low level of knowledge.					
Feeling that my confidence level matched my knowledge level encouraged me solve more questions.					
Feeling that my confidence level was different from my knowledge level encouraged me solve fewer questions.					
<b>Learner model charts: combined model</b>					
I can easily notice from the chart, when my confidence and my level of knowledge match completely.					
I can easily notice from the chart, when my confidence and my level of knowledge mostly match.					
I can easily notice from the chart, when my confidence and my level of knowledge match somewhat					
I can easily notice from the chart, when my confidence and my level of knowledge are completely different					
Being shown that there is a difference between my knowledge and confidence is helpful					
Seeing my model let me be more accurate in assessing my confidence level					
Being shown that there is a difference between my knowledge and confidence led me to solve more questions.					
Being shown that there is a difference between my knowledge and confidence led me to answer fewer questions.					
<b>OLMlets Expanded Charts</b>					
Seeing my confidence model beside my knowledge model is useful.					
Seeing my confidence model beside my knowledge model helped me to compare the two models.					
Seeing my confidence model beside my knowledge model helped me to be more accurate in assessing my confidence level.					
Seeing my confidence model beside my knowledge model encouraged me to solve more questions.					
Seeing my confidence model beside my knowledge model encouraged me to solve fewer questions.					
<b>Comments:</b> Is there anything else (good or bad) that you would like to tell us: ----- ----- -----					

## Appendix 8: Post-Questionnaire students responses (Chapter 5)

user_id	groups	Q1 Learning	Q2 Learning	Q3 Learning	Q4 Learning	Q5 Learning	Q6 Learning	Q1 Confidence	Q2 Confidence	Q3 Confidence	Q4 Confidence	Q5 Confidence	Q6 Confidence	Q7 Confidence
101572	1	5	4	3	3	4	3	4	5	4	4	5	3	3
100598	1	5	5	3	4	5	5	4	4	4	5	4	5	5
103426	1	5	4	3	4	4	3	4	2	4	4	2	3	3
101967	1	4	3	4	4	3	4	4	4	5	5	3	4	4
107446	1	4	4	3	5	4	3	2	4	4	3	3	3	3

107951	1	4	4	5	4	5	4	5	2	4	5	1	4	1
102480	1	4	3	3	4	3	3	4	4	3	4	2	3	3
100750	1	4	4	4	4	4	4	4	5	5	4	4	4	4
100674	2	5	4	3	4	4	5	3	5	4	5	5	4	4
93002	1	4	4	4	4	4	4	3	3	3	4	4	2	3
96028	2	4	3	3	4	4	4	4	4	4	4	4	4	3
96773	2	4	5	4	3	4	4	3	5	5	3	2	5	5
103389	2	5	4	3	3	4	5	5	3	4	4	5	4	4
107756	2	4	4	4	5	5	5	3	5	5	4	4	4	4
102171	2	4	5	4	4	4	4	3	4	5	4	4	3	4
101270	2	5	5	4	5		5	4	5	5	5	5	5	4

103447	3	4	4	4	5	4	4	4	4	4	4	4	4	4
102994	3	4	4	5	4	4	4	5	4	4	5	4	5	5
108300	3	4	4	3	3	4	4	4	3	5	4	4	5	4
99146	3	5	5	4	5	5	5	5	5	3	5	3	5	1
105932	3	5	5	5	5	5	5	5	5	5	5	5	5	5
110660	3	5	4	4	4	3	5	5	4	3	4	2	5	4

## Post-Questionnaire data: treatment conditions

Interpreting visualising alignment using combined model

user_id	OLMlets	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
96028	2	4	4	4	4	4	4	4	4	4	4
96773	2	4	4	5	5	2	2	5	5	5	2
103389	2	5	5	5	4	4	3	3	4	3	3
100674	2	5	4	5	4	4	5	4	4	4	2

107756	2	4	4	4	5	4	4	5	5	5	4
102171	2	4	4	4	5	3	5	4	4	5	5
101270	2	5	5	5	5	4	4	4	5	4	3
103447	3	4	5	4	4	4	4	2	4	4	2
102994	3	5	5	4	5	5	5	5	5	5	1
108300	3	4	5	4	4	4	4	4	4	4	4
99146	3	5	5	4	5	5	5	5	4	4	1
105932	3	5	5	5	5	5	5	5	5	5	5
110660	3	4	4	4	4	4	4	4	4	4	4



# Appendix 9: Pre-test (Chapter 5)

Sultan Qaboos University

Department of Mathematics

Introduction to complex Numbers – MATH4452

Student Name-----Student ID-----

Please answer the multiple-choice questions and the self-assessment as well

1. Solve the following problem.  $(4 + 3i) (2 + 2i)$
- a.  $2 - 14i$
  - b.  $6 + 5i$
  - c.  $2 + 14i$
  - d.  $2 + i$

How sure are you that your answer is correct?

- very unsure       unsure       sure       very sure.

2. Find the quotient  $\frac{(6 + 2i) - (1 + 3i)}{(-1 + i) - 2}$
- a.  $\frac{8}{5} - \frac{1}{5} i$
  - b.  $\frac{-8}{5} - \frac{1}{5} i$
  - c.  $\frac{8}{5} + \frac{1}{5} i$
  - d.  $\frac{-8}{5} + \frac{1}{5} i$

How sure are you that your answer is correct?

- very unsure       unsure       sure       very sure.

3. Determine the correct answer for the following expression.  $(i) (6i)$
- a.  $6i$

- b. 6
- c.  $6 - i$
- d.  $-6$

How sure are you that your answer is correct?

- very unsure       unsure       sure       very sure.

4. Which answer is correct for the following expression  $(1 + 3i)(-4 - 7i)(6 - 5i)$

- a.  $7 - 199i$
- b.  $3 + -9i$
- c.  $24 + 105i$
- d.  $4 + 30i$

How sure are you that your answer is correct?

- very unsure       unsure       sure       very sure.

5. If  $i^{4k} = 1$ , then  $i^{4k+2}$  and  $k$  is an integer

- a. 1
- b. -1
- c. -i
- d. 0

How sure are you that your answer is correct?

- very unsure       unsure       sure       very sure.

6. Evaluate the following complex number expression.  $(10)(8i)$

- a.  $2i$
- b.  $18i$
- c.  $-18$
- d.  $80i$

How sure are you that your answer is correct?

- very unsure       unsure       sure       very sure.

7. The conjugate of a complex number  $Z = a + ib$  is
- a.  $a^2 + b^2 i$
  - b.  $a - ib$
  - c.  $a^2 - b^2 i$
  - d.  $a + ib$

How sure are you that your answer is correct?

- very unsure       unsure       sure       very sure.

8. Evaluate the following expression.  $(1 + i)(2 + i)$
- a.  $2 + 2i$
  - b.  $2$
  - c.  $1 + 3i$
  - d.  $-1$

How sure are you that your answer is correct?

- very unsure       unsure       sure       very sure.

9. Multiply  $(1 + i)$  by its complex conjugate and simplify the answer
- a.  $2$
  - b.  $2+i$
  - c.  $1-i$
  - d.  $-2$

How sure are you that your answer is correct?

- very unsure       unsure       sure       very sure.

10. If  $Z = x + iy$ , then modulus of  $Z$ ,  $|Z| =$

a.  $\sqrt{(x^2 + y^2)}$

b.  $\sqrt{(x^2 - y^2)}$

c.  $\sqrt{(x + y)}$

d.  $\sqrt{(x - y)}$

How sure are you that your answer is correct?

very unsure

unsure

sure

very sure.

# Appendix 10: Post-test (Chapter 5)

Sultan Qaboos University

Department of Mathematics

Introduction to complex Numbers – MATH4452

Student Name-----Student ID-----

Please answer the multiple-choice questions and the self-assessment as well

1. Multiply  $(6 + 9i)$  by its complex conjugate and simplify your answer
  - a. 45
  - b. 117
  - c. 107
  - d. 54

Are you sure that your answer is correct?

very unsure    unsure    sure    very sure.

2. Subtract the following complex number  $(4 + 6i) - (12 - 3i)$ 
  - a.  $-8 + 3i$
  - b.  $-8 - 9i$
  - c.  $48 - 64i$
  - d.  $48 + 64i$

Are you sure that your answer is correct?

very unsure    unsure    sure    very sure.

3. Find the modulus and the argument of  $z$ , where  $z = 1 - \sqrt{3}i$

- a.  $2, \frac{5\pi}{3}$
- b.  $\frac{1}{2}, \frac{5\pi}{3}$
- c.  $2, \frac{\pi}{3}$
- d.  $\frac{1}{2}, \frac{\pi}{3}$

Are you sure that your answer is correct?

- very unsure    unsure    sure    very sure.

4. Simplify  $i^{24}$

- a. -1
- b. 1
- c. i
- d. -i

Are you sure that your answer is correct?

- very unsure    unsure    sure    very sure.

5. If  $z$  and  $w$  are two complex numbers, then

- a.  $\overline{z + w} = \overline{z} + \overline{w}$
- b.  $\overline{z + w} = \overline{z} \cdot \overline{w}$
- c.  $\overline{z + w} = \overline{z} - \overline{w}$
- d.  $\overline{z + w} = \overline{z} / \overline{w}$

Are you sure that your answer is correct?

- very unsure    unsure    sure    very sure.

6. If  $z = 9 - 3i$ , then  $|z|$  is

- a.  $\sqrt{99}$
- b.  $\sqrt{72}$
- c.  $\sqrt{12}$
- d.  $\sqrt{6}$

Are you sure that your answer is correct?

very unsure    unsure    sure    very sure.

7. The polar form of  $z = -2\sqrt{3} - 2i$ , where the argument is in the interval  $[0, 2\pi]$  is

- a.  $4cis\left(\frac{\pi}{6}\right)$
- b.  $4cis\left(\frac{7\pi}{6}\right)$
- c.  $4cis\left(\frac{\pi}{7}\right)$
- d.  $4cis\left(\frac{6\pi}{6}\right)$

Are you sure that your answer is correct?

very unsure    unsure    sure    very sure.

8. If  $z = 1 - \sqrt{3}i$ , then  $\text{Arg}(z)$  is equal

- a.  $\frac{\pi}{3}$
- b.  $2n\pi - \frac{\pi}{3}$ ,  $n$  is integer
- c.  $\frac{-\pi}{3}$
- d.  $2\pi$

Are you sure that your answer is correct?

very unsure    unsure    sure    very sure.

9. What is  $(12, 195^\circ)$  in Cartesian coordinates?

- a.  $(-11.59, -3.11)$
- b.  $(11.59, 3.11)$
- c.  $(3.11, 11.59)$
- d.  $(-11.59, -3.11)$

Are you sure that your answer is correct?

very unsure    unsure    sure    very sure.

10. If  $z = -1 + i$  then  $z$  expressed in polar form is

- a.  $\sqrt{2} cis \frac{3\pi}{4}$
- b.  $\sqrt{2} cis \frac{\pi}{4}$

c.  $\sqrt{2} \operatorname{cis} \frac{-\pi}{4}$

d.  $\operatorname{cis} \frac{-\pi}{4}$

Are you sure that your answer is correct?

very unsure    unsure    sure    very sure.



## Appendix 11: Students logs data (OLMlets) (Chapter 5)

user_id	pre-groups	OLMlets Use	post-groups	Total Questions	Total Views	Total Expanding model	Pre Confidence	Post Confidence	Confidence gain	Pre -test Performance	Post- test performance
102321	1	No_OLM	4	58	0	0	2.7	3.4	0.7	8	8
102166	2	No_OLM	4	27	0	0	2.4	1.6	-0.8	6	3

108357	1	No_OLM	4	26	0	0	2.6	2	-0.6	4	4
110660	3	No_OLM	4	26	0	0	2.56	3	0.44	7	5
107952	3	No_OLM	4	17	0	0	1.6	3.29	1.69	2	5
103236	2	No_OLM	4	9	0	0	2.5	3.22	0.72	6	7
100868	1	No_OLM	4	8	0	0	2.5	2.2	-0.3	4	5
101203	2	No_OLM	4	8	0	0	2.9	2.5	-0.4	10	4
101024	1	No_OLM	4	7	0	0	3.1	3.4	0.3	7	6
102234	3	No_OLM	4	5	0	0	2.3	4	1.7	5	9
107562	3	No_OLM	4	4	0	0	3.7	3.7	0	9	7
103238	1	No_OLM	4	0	0	0	3.4	3.1	-0.3	9	6
106022	1	No_OLM	4	0	0	0	3.6	3.1	-0.5	9	6

106808	1	No_OLM	4	0	0	0	1.78	2.6	0.82	4	6
107435	1	No_OLM	4	0	0	0	2.2	1.9	-0.3	5	5
111414	1	No_OLM	4	0	0	0	1.8	2.3	0.5	1	2
113333	1	No_OLM	4	0	0	0	3.8	3.9	0.1	9	7
94022	2	No_OLM	4	0	0	0	3.7	3.3	-0.4	9	6
94265	2	No_OLM	4	0	0	0	2.2	2	-0.2	4	1
97749	2	No_OLM	4	0	0	0	3.7	3.6	-0.1	8	9
100788	2	No_OLM	4	0	0	0	2.6	2.5	-0.1	8	7
100800	2	No_OLM	4	0	0	0	3.1	2.2	-0.9	7	3
103271	2	No_OLM	4	0	0	0	3.4	2.2	-1.2	8	7
106814	2	No_OLM	4	0	0	0	2.3	1.8	-0.5	5	4

108360	2	No_OLM	4	0	0	0	2	3.2	1.2	1	5
100711	3	No_OLM	4	0	0	0	3.5	2.6	-0.9	9	4
106359	3	No_OLM	4	0	0	0	2.8	2.6	-0.2	7	5
110928	3	No_OLM	4	0	0	0	2.9	2.9	0	5	5
113391	3	No_OLM	4	0	0	0	2.4	3.7	1.3	4	9
100598	1	OLM	1	26	1	0	3.8	3.78	-0.02	9	8
101967	1	OLM	1	26	2	0	3.2	3.7	0.5	9	7
103426	1	OLM	1	25	1	0	3.2	3.44	0.24	7	6
102480	1	OLM	1	18	1	0	3.1	3.1	0	5	7
107951	1	OLM	1	26	1	0	4	3.7	-0.3	10	8
100750	1	OLM	1	26	3	0	3.2	2.7	-0.5	9	6

101572	1	OLM	1	8	1	0	3.7	2.8	-0.9	10	6
107446	1	OLM	1	26	17	0	3.57	2.8	-0.77	7	3
93002	1	OLM	1	26	1	0	2.9	2.78	-0.12	7	5
101656	2	OLM	2	26	1	0	3.5	3	-0.5	8	5
96028	2	OLM	2	26	1	0	3	3.3	0.3	6	7
96773	2	OLM	2	32	3	0	3.5	3.56	0.06	9	5
102171	2	OLM	2	9	1	0	2.78	3.1	0.32	5	5
103389	2	OLM	2	57	4	0	3.6	3.5	-0.1	9	7
107855	2	OLM	2	7	1	0	3.4	3.4	0	10	7
107756	2	OLM	2	35	3	0	2.6	2.9	0.3	4	8
100674	2	OLM	2	82	17	0	4	3.89	-0.11	8	9

101270	2	OLM	2	64	3	0	2.8	3.8	1	9	8
97001	3	OLM	3	72	7	1	2.7	3.44	0.74	10	5
100819	3	OLM	3	58	32	17	3.1	3.8	0.7	5	7
99146	3	OLM	3	56	3	3	3.7	3.3	-0.4	9	4
103447	3	OLM	3	14	5	0	2.1	4	1.9	5	9
102994	3	OLM	3	64	7	7	4	4	0	9	9
108300	3	OLM	3	58	3	3	1.5	3.3	1.8	4	6
105932	3	OLM	3	93	5	3	1.9	4	2.1	0	6

## Appendix 12: Students logs data (OLMlets) (Chapter 6)

user_id	user_id	group	study ability level	Total Questions	Total views	Total number of expanding the model	Set1 Bias	Set 1 confidence	Set1 performance	Set1 questions answered	Set1 views	Set2 Bias	Set2 confidence	Set 2 Performance	Set2 questions answered	Set 2 views	No.attempts Set 1	No.attempts Set 2
S1	248	1	Low	150	46	-	0.43	3.84	0.52	50	18	0.32	4	0.68	50	8	2	2
S8	256	1	Low	275	67	-	0.17	2.6	0.36	75	4	-0.01	3.04	0.68	125	46	3	0.8
S16	264	1	Low	75	71	-	0.1	3	0.56	25	23	-0.02	3	0.68	25	24	1	0.8
S20	270	1	Low	245	210	-	-0.01	2.8	0.6	100	95	-0.29	2.8	0.88	100	75	4	1
S26	276	1	Low	313	250	-	0.01	2.84	0.6	177	123	-0.1	3	0.76	76	75	7.1	2
S28	278	1	Low	50	2	-	0.09	2.84	0.52	25	1	-0.15	2.6	0.68	25	1	1	
S36	286	1	Low	200	150	-	0.05	2.48	0.49	100	61	-0.06	3	0.72	50	50	4	1
S2	249	3	Low	188	147	2	-0.24	1.96	0.56	50	33	0.17	3.92	0.8	78	58	2	5
S9	257	3	Low	200	9	1	-0.23	2.12	0.6	125	6	-0.33	1.68	0.56	50	2	5	1.56
S11	259	3	Low	125	94	82	0.2	3.4	0.6	100	69	0.15	3.72	0.76	25	25	4	3.4
S15	263	3	Low	150	150	99	-0.18	1.92	0.48	50	50	-0.02	3	0.68	50	50	2	3.04

S18	267	3	Low	285	188	0	-0.11	2.72	0.57	100	72	-0.14	2.76	0.72	125	75	4	1
S27	277	3	Low	74	38	0	-0.04	2.92	0.68	29	1	-0.07	2.84	0.68	25	17	1.2	1.88
S29	279	3	Low	55	6	2	-0.15	2.6	0.68	26	2	-0.15	3.22	0.89	9	2	1	1
S33	283	3	Low	95	86	2	-0.03	2.72	0.6	50	43	-0.21	2.8	0.8	25	25	2	0.96
S35	285	3	Low	174	94	3	0.36	3.52	0.48	75	45	0.17	3.92	0.8	24	14	3	2
S39	291	3	Low	191	165	0	-0.05	2.43	0.52	75	68	-0.18	2.4	0.64	75	68	3	3
S40	293	3	Low	32	6	6	-0.2	1.84	0.48	32	6						1.3	
S4	252	1	High	51	37	-	0.23	3.84	0.72	31	22	0.25	3.55	0.6	20	15	1.2	3.12
S10	258	1	High	52	52	-	-0.11	2.96	0.76	31	31	-0.24	3	0.9	20	20	1.2	0.92
S12	260	1	High	125	105	-	0.29	3.8	0.64	60	47	0.36	4	0.64	40	33	2.4	5
S14	262	1	High	24	12	-	0.22	4	0.78	24	12						1	2
S24	274	1	High	145	93	-	-0.03	2.96	0.68	25	8	-0.24	2.56	0.76	85	61	1	1.6
S30	280	1	High	75	73	-	0.14	3.36	0.64	25	24	-0.04	3.28	0.8	25	24	1	2
S34	284	1	High	50	2	-	0.01	3.08	0.68	25	1	-0.08	3.04	0.76	25	1	1	1
S7	255	3	High	67	29	0	-0.08	3.09	0.77	22	9	-0.08	3.13	0.78	23	10	0.9	4
S13	261	3	High	105	57	5	0.05	3.68	0.84	25	17	0.13	3.8	0.8	50	24	1	2
S17	265	3	High	75	19	0	0.02	3.72	0.88	25	1	-0.14	3.48	0.96	25	1	1	1
S21	271	3	High	163	131	0	0.01	2.96	0.64	57	45	0.02	3	0.64	39	37	2.3	0.36
S25	275	3	High	95	4	3	-0.2	2.56	0.72	25	1	-0.31	2.24	0.72	50	2	1	1
S31	281	3	High	117	98	1	-0.06	3.24	0.8	50	39	-0.09	2.92	0.72	47	41	2	1
S37	289	3	High	376	302	1	0.2	3.76	0.72	100	77	0.47	3.96	0.52	134	100	4	5.36



# Appendix 13: Semi-structured Interview (Chapters 6 & 7)

1. What was your target while using OLMlets?
2. How did you use the System? (alternative: Please describe how you used OLMlets?)
3. Why did you choose to use OLMlets?
4. You were asked to assess your confidence on each question. What did you think of this?
  - a. Is this the same way you felt last term (alternative: in the first experiment)? If so, how?
5. What did you think of how OLMlets gave you information about your confidence and performance (using skill meters)?
  - a. How did this information make you feel?
6. How did you use the information that was given to you about your confidence and your performance in solving the system's questions?
  - a. What did you do as a result of seeing this information?
  - b. How is this different from before?
7. How good were you at determining what you know before using OLMlets?
  - a. How good do you think you are at assessing your knowledge now?
8. How do you feel your performance changed in first semester?

9. How do you feel your performance changed this past term?
10. Which learner model visualisation you prefer to use when studying?
11. Do you have anything else you want to share about your experience in using the OLMlets system?

## Appendix 14: Semi-structured interview (Students quotes classified by themes)

Student Id: 276 (S26) Low Performing student in both study 1 & study 2 Set 1 (Original skill meter /Expandable model)	
Proposition	Category
I used [OLMlets] mainly for exam preparation	Exam preparation
going in the exam	Exam preparation
It mainly be some way of supplementary revision	Exam preparation
just to make sure that I need exactly what I want I needed to know	
If everything was right then, fine I won't be worried about it	understanding of topic
Is more I just I wanted to know where over I was getting the questions right or wrong	understanding of topic
So, if I got one or two wrong I will not aware about it, but if I got the whole chunk wrong, then really, I need to worry about it.	understanding of topic
so <b>first I start going to revise whatever topic it was in OLMlets covered</b>	Revision strategy
and then after I felt I revised well, I go to the OMLlets, answer couple of questions seeing what the result was	Revision strategy
if it was then I sort of go back and use as a way of saying that I need to review this area.	Revision strategy

<b><i>If I got staff wrong then I will go back and I will read all the stuff and if I got the stuff wright then I will not worry about it.</i></b>	Revision strategy
Honestly, it kind of got a bit annoying after a while	Confidence in understanding (-)
it didn't feel like it made much difference.	Confidence in understanding (-)
rather than where I feel confident or not	Confidence in understanding (-)
In terms of confidence, it didn't make that much of the difference, kind even if got high or low confidence.	Confidence in understanding (-)
honestly it more or less helped about the same amount	Confidence in understanding (-)
So, because at the same point was I can feel quite confident in most of the questions,	Confidence in understanding
even putting things I don't know and I feel not confident in this question	Confidence in understanding
Usually if it was high confidence but I know I got more of them wrong, I kind use that to go to do over to use some practical approach to revise in that sort of going back and coding	Confidence in understanding
I think defiantly to see the performance is quite helpful to seeing that, sort of graphic to know.	Visualisation Preference
I think the [expandable model] was better	Visualisation Preference
I probably say the [ expandable model] because you have sort of it shows you how much you certain or uncertain you are.	Visualisation Preference

<b>Student Id: 274 (S24) High Performing student on both study 1 &amp; 2</b>	
<b>Set 4 (Original skill meter /text-based model)</b>	
<b>Proposition</b>	<b>Category</b>
I can know my weakness like in my knowledge	Understanding a topic
I did it as a test kind of without knowing what I got.	Revision strategy
I didn't do any book work before OLMlets	Revision strategy
I think there were too many options kind of	Self-assessment (-)
I think it was really good actually	Confidence in understanding (+)
I thought I was confident with that particular one but I wasn't. When it did, it helped a lot.	Confidence in understanding (+)
I wanted to be confident and correct.	Confidence in understanding (+)
I felt much more confident after OLMlets.	Confidence in understanding (+)
the confident bar side you can know how you understand the question,	Confidence in understanding (+)
I prefer the skill meters'	Visualisation preference

<b>Student Id: 262 (S14)</b>		<b>High Performing student on both study 1 and study 2</b>	
		<b>Set 2 (Original skill meter/Text-based model)</b>	
<b>Proposition</b>		<b>Category</b>	
sort of a revision,		preparation	
was good for that because it actually modelled what we actually get in the exam.		preparation	
more focused on wither what I thought was the right answer and what was not the right answer.		Understanding a topic	
sort of repeating the questions over and over		Revision strategy	
two windows side by side,		Viewing strategy	
I didn't used it to find how confident on answering questions		Confidence in understanding (-)	
more focused to just to be green		Confidence in understanding (-)	
I think it was (low, moderate, high, and very high), which I guess is an indication of how confident you are. But it is not really good measure how well you are in terms of doing answering the questions		Representing confidence in text-based (-)	
Sometimes you will answer one question and it will go tiny bit or you will get one wrong and it will crash the way down		Understanding OLMlets learner model processing (-)	
based on recent performance rather than the overall performance		Understanding OLMlets learner model processing (+)	

I didn't realize that there was an option [Expanding the model]	Knowing OLMlets features (-)
I said that I am very certain through all the questions.	Confidence judgment (-)
I much prefer the visual one (skill meter),	Visualisation preference
but the confidence I think the text based is more representative	Preference for confidence visualisation

<b>Student Id: 271 (S21)      High Performing student in both study 1 &amp; study 2</b> <b>Set 3 (Expandable model/ Text-based model)</b>	
<b>Proposition</b>	<b>Category</b>
to test my strength and weakness	Understanding a topic
I go to the model and check the graphs and see if I got it right or wrong. Then if I got it wrong, I will go back to my notes	Understanding a topic
revise it first and then I will go through each the sections and then I will answer all the questions	Revision strategy
kept solving them until I got all of them correct	Revision strategy
I will go through the questions individually and then check the model.	Viewing strategy
it was helpful to get a few what the class test will look like	Motivation to use OLMlets
it was quite difficult <b><i>because I didn't want to answer very sure in case I got wrong or completely not sure.</i></b>	Confidence judgment (-)
I don't know really how good in anything until I actually try answer question then you can know what you know and what you don't know,	Confidence judgment (-)
it is a lot higher [accuracy in knowledge after OLMlets] once I gone through it.	Confidence in understanding



basically it will not make a huge difference	Confidence in understanding (-)
I liked the visual one, even with a slight change you can see the difference,	Visualisation preference (+)
To have it less opaque if you less confident or opaquer if you are more confident will be quite good.	Visualisation preference (+)
just like checking my performance was easier to use the first model	Visualisation preference (+)
text based you obviously need more time, will not update after each question so it can be moderate for three question	Visualisation preference (-)

<b>Student Id: 252 (S4) High Performing student on both study 1 &amp; study 2</b>	
<b>Set 1 (Original skill meter/Expandable model)</b>	
<b>Proposition</b>	<b>Category</b>
mainly using it for revision tool	Exam preparation
OLMlets helped me to cover wide range of topics	Exam preparation
I see question then I look first I look to the model	Viewing strategy
got it right and putting not sure but not very sure that will change the green to something lighter, I don't think there is a need for.	Confidence in understanding (-)
I didn't pay much attention to it	Confidence in understanding (-)
as it was green I was fine	Confidence in understanding (-)
I didn't pay much attention to the confidence level to be honest (Expandable model)	Confidence in understanding (-)
If I saw the answers from my lecture notes I click in the answer right there, but if I did not see it I click the answer, any answer I thought it was right,	OLMlets miss use
the ones which I wasn't sure, I already knew that I am not sure about them	Confidence judgment (+)
the first one is more straight forward (original skill meter)	Preference visualisation (+)

prefer the graphical one [compared to text-based]	Preference visualisation (+)
the first one is more straight forward (original skill meter)	Preference visualisation (+)

<b>Student Id: 260 (S12) High Performing student in study 1 and low in study 2</b>	
<b>Set 1 (original skill meter/Expandable model)</b>	
<b>Proposition</b>	<b>Category</b>
if I wasn't sure <i>I used OLMlets before I revise for one topic and then after I revised the whole booklet.</i>	Revision strategy
To test myself and it kind helped to learn as well	Understanding a topic
open the tap with the model, but then I realized that I can just open two tabs	Viewing strategy
we had it OLMlets but without it [confidence assessments] . It was much better,	Confidence in understanding (-)
wouldn't pay much attention on it. [ confidence]	Confidence in understanding (-)
So after I revise the booklet I was sure in most of the questions.	Confidence in understanding (+)
I will say 65 [before] , I think 75 [after]	Confidence judgment (+)
I will choose unsure or very confident (very sure) and I couldn't tell much difference in colours may be because <i>most of the time I will choose very confident or just in the middle.</i> So only two of them.	Interpreting opacity of the colour (-)
think it was easy to see the transparent green bar	Interpreting opacity of the colour (+)
know I guess I feel much happier if I seen the transparent green	Interpreting opacity of the colour (+)
I wouldn't want that [text-based]	Visualisation Preference (-)

<b>Student Id : 280 (S30) High Performing student in study 1 and Low in study 2 Set 1 (Original skill meter /Expandable model)</b>	
<b>Proposition</b>	<b>Category</b>
I used it to prepare for the class test	Exam preparation
I wanted to test my knowledge	Understanding a topic
I sort put in the side as I do the questions I open another tab so I can check what I got wrong.	Viewing strategy
in some cases, I thought it is useful,	Confidence in understanding (+)
Like sometimes I feel some rush I need to go through and then I need to keep clicking.	Confidence in understanding (-)
I was using it, the confidence thing been compulsory was a little bit annoying.	Confidence in understanding (-)
Highlight a lot potential problems that I didn't pickup from before.	Confidence in understanding (-)
<b><i>I think most of the times I will click very sure, just click that.</i></b>	Confidence in understanding (-)
Just in a few cases I choose the actual confidence bar.	Confidence judgment (-)
the times that I used it honestly then It did, it did give a bit of idea	Confidence in understanding (+)
<b><i>If it was more transparent then I will go over the questions a bit more until I found confident</i></b>	Confidence in understanding (+)

that I have more time to fully prepare for it, then the knowledge and the confidence is good	Confidence in understanding (+)
Highlight a lot potential problems that I didn't pickup from before.	Confidence in understanding (+)
Definitely helped [OLMlets use] a lot to bring up.	OLMlets use (+)
Say the confidence one, yes.	Visualisation preference

<b>Student Id: 264 (S16)                      Low Performing student in both study 1 &amp; study 2</b>	
<b>Set 1 (Original skill meter /Expandable model)</b>	
<b>Proposition</b>	<b>Category</b>
I try to learn the content and then I will go to OLMlets and do the relevant questions and then obviously I look at the model.	Revisions strategy
If I didn't understand anything or It showed that I didn't understand a certain topic I will go back and then I will re do the OLMlets.	Revisions strategy
without answering a question you never really know wither you understand it, so it 100 % definitely helped a lot	Understanding a topic
after each question [views]	Viewing strategy
I can tell myself If I am confident in my answer I didn't need the model to help me in that.	Self-Beliefs of knowledge monitoring skills
I was putting it [confidence judgment] every time what I felt to use it.	Confidence judgment (+)
very satisfying the ones you knew something with confident it was the right colour it was nice dark green	Confidence in understanding (+)
It helped to understand what I am thinking but it didn't help to learn	Confidence in understanding (+)
was more transparent I was definitely more likely to go back and do it again	Confidence in understanding (+)

during revision before that, it is more helpful to have the opacity so it helped me engage how well I knew something not just I know it.	Confidence in understanding (+)
I had quiet a good knowledge base anyway so the OLMlets wasn't as helpful but with UML I had much less knowledge	Student Domain knowledge
definitely the opaque one	visualisation preference
better to have the second one	visualisation preference
I liked the shading it was very clear to see	Interpreting opacity visualisation (+)



Student Id: 289 (S37) High Performing student in study 1 and low in study 2 Set 4 (Expandable model/Expandable model)	
Proposition	Category
so I <i>can practise for the class test</i>	Exam Preparation
answering the questions and checking if I got it right,	Understanding a topic (+)
I wanted at the end to answer the whole of the questions right and have a solid green bar	Understanding a topic (+)
having instant feedback helped a lot to say oh I need to work, look at the lecture notes in this area and maybe you don't want to look this other area.	Understanding a topic (+)
<b><i>doing the OLMlets really helped with like gaining new knowledge and consolidating the old knowledge</i></b> , yes, the feedback was good	Understanding a topic (+)
It is hard to judge confidence in like four categories, generally I answered either very confident or not sure	Confidence judgment (-)
I thought, actually before OLMlets I thought I knew a lot but then doing the OLMlets like woo, these areas I don't know	Confidence judgment (+)
I didn't really see the point of the confidence	Confidence in understanding (-)

I don't think I made attention to the transparency as much green or red	Confidence in understanding (-)
after a while did I got the questions right or I got it wrong.	Confidence in understanding (-)
Initially I would do it more slowly and can think about the questions if I am sure or not sure	Confidence in understanding (+)
before the exam I will use the confidence it is more helpful,	Confidence in understanding (+)
make sure it was green like I knew it solid green	Confidence in understanding (+)
confidence in knowing that I can do these questions I am more confident in these topics now	Confidence in understanding (+)
did well in java anyway I knew a lot of java anyway. UML was I didn't know any UML	Domain Knowledge
The green staff was really obvious but the transparency wasn't as obvious.	Interpreting opacity levels (-)
the transparency wasn't as certain as this one.	Interpreting opacity levels (-)
would be more clear. Having two bars instead of one.	visualisation preference
I didn't want to have that text. ..because the bars are so easy you just look at them	visualisation preference
I am happy with this one [Expandable] I think it work well and I don't need to change it.	visualisation preference

<b>Student Id 257 (S9)</b> <b>Low Performing student in both study 1 and study 2</b> <b>Set 4 (Expandable model/Expandable model)</b>	
<b>Proposition</b>	<b>Category</b>
<i>I was using it to revise for tests.</i>	Exam Preparation
I think it made me feel more prepared for the test,	Exam Preparation
I then have to do the test like six times, to figure out which ones I got right and which one I got wrong	Revision strategy
I kind used the OLMlets before I did revision	Revision strategy
I did OLMlets first and then last, so I did the <b><i>OLMlets to see where I roughly I was and then normally the green bar were quiet low, and then I did some revision on kind of what the questions I asked and staff and then I go back and do it</i></b> and then do it again.	Revision strategy
then went back over and did all again like kept doing it until I got 100%,	Understanding a topic
for the java I did the whole thing then looked at the grade like all of the questions	Viewing Strategy
with UML I did it one at time so I did one question and went on so if I got it right or wrong.	Viewing Strategy
I would had the ones I would get right...I would have 50 % I will be yes definitely I know that and 50% would be like a guess. So, it was quite good	Confidence in understanding (+)

I think for me I was just seeing if the answers are right or wrong I didn't care if I am confident or not	Confidence in understanding (-)
it didn't matter if it was like transparent or not it was just as long it is green some form of green it was fine.	Confidence in understanding (-)
because you cant see the answers so the green bar will go down again so that will stress me out	Confidence in understanding (-)
used it like I did clicked confident or not confident.. it didn't helped me I would be fine with either way.	Confidence in understanding (-)
I think the idea that you learn to answer the questions and you don't learn the answers of that questions because that will not going help you when it is different in the exam.	Understanding OLMlets Aim
The text based, no I liked the bar.. it is easy to see 70 on a bar than it is in text.	Visualisation Preference
I wouldn't like as a text, I like the colours	Visualisation Preference

<b>Student Id : 283 (S33)</b> <b>Low Performing student in study 1 and high in study 2</b> <b>Set 3 (Expandable model /Text-based model)</b>	
<b>Proposition</b>	<b>Category</b>
helpful for revision	Exam Preparation
use it for revision	Exam Preparation
so I can go back and revise it.	Revision Strategy
so <b><i>I can understand what I know and what I didn't know</i></b>	Understanding a topic
if it is low then I go and revise it and if it is quite high then I am satisfied.	Understanding a topic
helped a bit in a sense that you get to understand	Understanding a topic
after one question I view it	Viewing Strategy
obviously figuring good marks and then you seen it from the feedback, your confidence raises	Confidence in understanding (+)
I think I picked the high and low more or less. Because after some time your confidence gets higher and you see it as well.	Confidence in understanding (+)
Then if you really get low and then your confidence drops obviously and then just fluctuates.	Confidence in understanding (+)
you have based on the answers you get and then you get to <b><i>see your knowledge and then you get to see wither the system think of your knowledge and then that's really helps</i></b>	Confidence in understanding (+)
sometimes you just guess it and then you know your confidence is low but you just guess it and then the system thinks you high knowledge but then it helps you	Confidence in understanding (+)

I want my confidence to be high but the knowledge I wanted to be very high.	Confidence in understanding (+)
I am going to answer more questions if I had that way [knowledge: high & confidence: low]	Confidence in understanding (+)
when you do the OLMlets you get better answers, then you say ok, this actually helps	Confidence in understanding (+)
I think the confidence level increases for the exams when you do all the revisions	Confidence in understanding (+)
I think it affects your confidence because you want to see move from moderate to high [text-based]	Confidence in understanding (-)
fairly good at doing that [confidence judgment]	Confidence judgment (+)
before using OLMlets I would say 80% but then after using it like 65 70%. because you get to see that you ok you actually misunderstood something or had a misconception	Self-Beliefs knowledge monitoring skills (-)
<b><i>the text based you don't really get to engage it well</i></b>	Visualisation Preference
when you are getting questions you see the little increase in the bar.. you don't get to see [text] just happen to be after certain questions just change to moderate.	Visualisation Preference
<b><i>I think a mixture of both is something that I will prefer.</i></b>	Visualisation Preference
I will prefer the text based.	Visualisation Preference

<b>Student Id: 267 (S18)                      Low Performing student in Study 1 and high in Study 2</b> <b>Set 3 (Expandable model /text-based model)</b>	
<b>Proposition</b>	<b>Category</b>
good for revising for the test	Exam Preparation
<i>helped me to get the knowledge like I was missing on</i>	Understanding a topic
<i>which ever questions I was not confident I was noted down and then I will do another go until I get right like</i>	Revision Strategy
Sometimes to do like one questions when I was felt that I am not sure about that questions.	Viewing strategy
I was clicking it random,	Confidence judgment (-)
the first go I will just go through I scan through the questions.. the second time [seriously] , like <i>I didn't do the questions one time I did it multiple times as I am going to the topic I was using the self-confidence more seriously.</i> the first round was just scan random	Confidence judgment (-)
for the second one I actually used it more seriously because it was more like texture one.	Confidence judgment (+)

confidence I felt my confidence was getting better	Confidence in understanding
mainly focused on the system's answer because I told you that I was to note down when I feel unconfident	Confidence in understanding (-)
I prefer the first one the java one the bars not the text one. It just graphical.	Visualisation Preference
it was quiet like difficult for the text one to for me to know the right answer sometimes it just remains same level, so I don't know if it was right or wrong.	Visualisation Preference
<b><i>because I got the texture one so I was a bit of taking more time to get the right answer</i></b>	Visualisation Preference
<b><i>I personally prefer the first one [skill meter]</i></b>	Visualisation Preference



<b>Student Id : 270 (S20)</b> <b>Low Performing student in both study 1 &amp; study 2</b> <b>Set 2 (original skill meter/ text-based model)</b>	
<b>Proposition</b>	<b>Category</b>
to come familiar with the style of questions that are going to be in the test	Exam preparation
<i>if I it was wrong I look for the answers straight way I just stop doing them and look at the topic in the book and then do the questions again</i>	Revision Strategy
so <i>to know how much I know</i> I have to test myself. So, the only test I did were from tutorials point website.. then I will go to OLMlets.	Revision Strategy
first time round I do it with the book answer the questions and then the second time round I will not have the book	Revision Strategy
<i>I did it the first few times [accurate confidence judgment]</i>	Confidence in understanding (+)
if the system was high and the confidence was low I will keep doing the questions	Confidence in understanding (+)
when I answered the confidence questions correctly then yes, they will help	Confidence in understanding (+)

afterwards I just wanted to do the questions so I didn't actually pay tension to that	Confidence in understanding (-)
<i>going from high to moderate obviously the last questions wrong</i>	Understanding Text-based LM
<i>I think I prefer the skill meter.</i>	Visualisation Preference
I was going through I just I didn't look at both but I used the expandable model just to see	Visualisation Preference

## Appendix 15: Students logs data (OLMlets) (Chapter 7)

id	conditions	ability level	post-groups	Total Questions	Total views	Total expanding models	Set 1 questions	Set 1 bias	Set 1 Confidence	Set 1 performance	Set 1 OLM views	Set 2 questions	Set 2 bias	Set 2 confidence	Set 2 Performance	Set 2 OLM views
S3	text-based	Low	1	76	5	-	50	0.08	3.04	0.6	4	25	0	4	1	1
S6	text-based	Low	1	37	1	-	25	0.41	3.68	0.48	1					
S7	text-based	Low	1	21	6	-	13	0.07	3.08	0.62	3	7	0.12	2.88	0.5	2
S10	text-based	Low	1	52	49	-	50	0.13	2.84	0.48	47	2				2
S25	text-based	Low	1	107	22	-	56	-0.1	2.64	0.64	20	25	-0.35	2.48	0.84	1
S20	text-based	Low	1	25	6	-	25	-0.1	2.4	0.56	6					
S39	text-based	Low	1	224	205	-	105	0.23	3.72	0.68	100	36	0.36	4	0.64	33
S22	Expandable	Low	2	119	44	1	25	0.2	3.4	0.6	1	44	0.14	3.48	0.68	19
S26	Expandable	Low	2	336	301	0	125	0.28	3.76	0.64	125	125	0.21	3.92	0.76	117
S30	Expandable	Low	2	52	50	0	21	-0.09	2.56	0.62	21	19	0.14	3.16	0.58	18
S37	Expandable	Low	2	324	287	0	158	0.53	3.92	0.44	145	90	0.2	4	0.8	77
S5	Expandable	Low	2	213	181	90	100	-0.16	2.04	0.5	73	48	-0.64	1	0.64	45
S9	Expandable	Low	2	140	135	0	65	0.44	4	0.56	63	25	0.2	4	0.8	25

S12	Expandable	Low	2	294	227	0	94	0.02	3	0.64	77	50	0.16	3.76	0.76	48
S16	Expandable	Low	2	135	132	0	18	-0.22	2.4	0.68	18	46	-0.16	2.44	0.64	45
S19	Expandable	Low	2	140	100	4	75	0.41	2.96	0.24	47	50	0.25	2.96	0.4	39
S23	Expandable	Low	2	54	50	0	25	-0.13	2.8	0.72	24	25	-0.18	3	0.84	22
S35	Expandable	Low	2	56	51	0	31	0.1	2.76	0.48	30					
S14	text-based	High	3	62	59	-	26	0.16	3.88	0.8	22	25	0.12	4	0.88	25
S15	text-based	High	3	25	25	-	25	-0.22	2.64	0.76	25					
S18	text-based	High	3	225	164	-	100	-0.36	2.2	0.76	87	50	-0.07	3.68	0.96	22
S21	text-based	High	3	100	83	-	25	-0.2	3.16	0.92	21	25	-0.3	3.12	1	19
S29	text-based	High	3	25	1	-	25	-0.67	2	1	1					
S33	text-based	High	3	75	52	-	25	-0.12	3.16	0.84	4	25	-0.08	2.2	0.48	25
S24	text-based	High	3	96	79	-	25	0.04	3.52	0.8	25	25	0.15	3.96	0.84	19
S32	text-based	High	3	100	7	-	25	-0.26	2.64	0.8	1	25	-0.4	2.68	0.96	3
S36	text-based	High	3	307	293	-	117	-0.1	3	0.76	81	90	0.22	3	0.44	116
S38	text-based	High	3	18	3	-	18	0.13	3.72	0.78	3					
S34	Expandable	High	4	75	69	0	25	0.16	3.76	0.76	21	25	0.32	4	0.68	25
S40	Expandable	High	4	133	77	0	25	-0.92	1	0.92	1	35	-0.48	2.32	0.92	9
S1	Expandable	High	4	208	102	2	54	0.2	4	0.8	33	44	0.04	4	0.96	19
S2	Expandable	High	4	197	188	0	59	-0.32	2.44	0.8	54	65	0.24	4	0.76	66
S4	Expandable	High	4	125	32	0	50	0	4	1	12	50	0.04	4	0.96	13
S8	Expandable	High	4	109	53	0	25	0.19	3.96	0.8	1	25	0.44	4	0.56	1
S13	Expandable	High	4	51	46	5	36	-0.03	3.32	0.8	31	3	0.11	3.33	0.67	3
S17	Expandable	High	4	57	50	0	9	0.04	3.78	0.89	7	24	-0.13	3.25	0.88	22

S27	Expandable	High	4	216	202	47	75	-0.11	2.96	0.76	75	50	-0.12	2.92	0.76	46
S31	Expandable	High	4	137	115	0	50	-0.26	3	0.92	44	32	-0.3	3	0.96	29