Learners' Self-Assessment and Metacognition when Using an Open Learner Model with Drill Down

MATTHEW DAVID JOHNSON

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

Department of Electronic, Electrical and
Systems Engineering
School of Engineering
College of Engineering and Physical Sciences
The University of Birmingham
July 2017

ACKNOWLEDGEMENTS

With thanks to Prof. Chris Baber for his guidance in the completion of this thesis and to Dr. Susan Bull for supervision during the earlier parts of the research.

Thanks also to Robert Gibson, Julia Johnson, David Johnson and countless friends and family who have encouraged the completion of the work and provided support and provided an environment that has fostered its momentum.

Special thanks also to Alice Kerly, Mary Winkles, Samantha McCauley and Dana Vasiljevic for being an executive sounding board during my time at the university.

Further thanks also to the FP7 projects NEXT-TELL and LEA's BOX, and to the University of Birmingham for the employment opportunities, which have taken place alongside the PhD research. Thanks also to the EPSRC for the 2 years 6 months of funding at the start of the research, and to Montvieux Ltd. for allowing some time for the preparation of the final draft.

ABSTRACT

Metacognition is ‘thinking on thinking’. It is important to educational practices for learners/teachers, and in activities such as formative-assessment and self-directed learning. The ability to perform metacognition is not innate and requires fostering, and self-assessment contributes to this. Literature suggests proven practices for promoting metacognitive opportunities and ongoing enquiry about how technology best supports these. This thesis considers an open learner model (OLM) with a drill-down approach as a method to investigate support for metacognition and self-assessment.

Measuring aspects of metacognition without unduly influencing it is challenging. Direct measures (e.g. learners ‘thinking-aloud’) could distort/disrupt/encourage/effect metacognition. The thesis develops methods to evaluate aspects of metacognition without directly affecting it, relevant to future learning-analytics research/OLM design. It proposes a technology specification/implementation for supporting metacognition research and highlights the relevance of using a drill-down approach.

Using measures that correspond to post-hoc learner accounts, this thesis identifies a baseline of student activity that is consistent with important regulation of cognition tasks and students’ specific interest in problems. Whilst this does not always influence self-assessment accuracy, students indicating their self-assessment ability can be used as a proxy measure to identify those who will improve. Evidence supports claims that OLMs remain relevant in metacognition research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

SA	Self-Assessment
ITS	Intelligent Tutoring System
LM	Learner Model
OLM	Open Learner Model / Open Learner Modelling
SD	Standard Deviation
SDR	Self-Directed Learning
STEM	Science, Technology, Engineering and Mathematics
EESE	School of Electronic, Electrical and Systems Engineering, University of Birmingham

The terms “student” and “learner” may be used interchangeably. These are one and the same stakeholder type.

LIST OF PUBLICATIONS

Book Chapters

Johnson, M.D., Bull, S. (2016). Designing for Visualisation of Formative Information on Learning, in Reimann, P., Bull, S., Kickmeier-Rust, M., Vatrapu, R., Wasson, B. (Eds.) Measuring and Visualizing Learning in the Information-Rich Classroom. Routledge: 2016.

Bull, S., **Johnson, M.D.**, Masci, D., Biel, C. (2016). Integrating and Visualising Diagnostic Information for the Benefit of Learning, in Reimann, P., Bull, S., Kickmeier-Rust, M., Vatrapu, R., Wasson, B. (Eds.) Measuring and Visualizing Learning in the Information-Rich Classroom. Routledge: 2016.

Biel, C., Ciernak, G., **Johnson, M.D.**, Bull, S., Hesse, F.W (2016). Influencing Cognitive Density and Enhancing Classroom Orchestration, in Reimann, P., Bull, S., Kickmeier-Rust, M., Vatrapu, R., Wasson, B. (Eds.) Measuring and Visualizing Learning in the Information-Rich Classroom. Routledge: 2016.

Conference Proceedings

Ginon, B., **Johnson, M.D.**, Turker, A., Kickmeier-Rust, M.D. (2016). Helping Teachers to Help Students using an Open Learner Model. To appear in the proceedings of European Conference on Technology Enhanced Learning (EC-TEL) 2016.

Ginon, B., **Johnson, M.D.**, Boscolo, C., Bull, S. (2016). Supporting Students of Italian as a Foreign Language with Learner Model Persuasion: An Exploratory Study. Proceedings of Intelligent Tutoring Systems (ITS 2016), Zagreb, Croatia.

Bull, S., Ginon, B., Boscolo, C., **Johnson, M.D.** (2016). Introduction of Learning Visualisations and Metacognitive Support in a Persuadable Open Learner Model. Proceedings of Learning, Analytics and Knowledge (LAK 2016), Edinburgh, UK.

Bull, S., **Johnson, M.D.**, Demmans Epp, C., Masci, D., Alotaibi, M. & Girard, S. (2014). Formative Assessment and Meaningful Learning Analytics, Proceedings of ICALT 2014.

Bull, S., **Johnson, M.D.**, Alotaibi, M., Byrne, W. & Cierniak, G. (2013). Visualising Multiple Data Sources in an Independent Open Learner Model, in H.C. Lane & K. Yacef (eds), Artificial Intelligence in Education.

Bull, S., Kickmeier-Rust, M., Vatrapu, R., **Johnson, M.D.**, Hammermueller, K., Byrne, W., Hernandez-Munoz, L., Giorgini, F., Meissl-Egghart, G. (2013). Learning, Learning Analytics, Activity Visualisation and Open Learner Model: Confusing?, Eighth European Conference on Technology Enhanced Learning, Paphos, Cyprus, September 2013.

Johnson, M.D., Bull, S., Wasson, B., Hansen, C., Cierniak, G., Debus, K., Biel, C. (2013). Initial Use of a Flexible Open Learner Model, in Wong, L.-H. et al. (Eds.) (2013). Proceedings of the 21st International Conference on Computers in Education. Indonesia: Asia-Pacific Society for Computers in Education.

Johnson, M.D., Cierniak, G., Hansen, C., Bull, S., Wasson, B., Biel, C., Debus, K. (2013).Teacher Approaches to Adopting a Competency Based Open Learner Model, in Wong, L.-H. et al. (Eds.) (2013). Proceedings of the 21st International Conference on Computers in Education. Indonesia: Asia-Pacific Society for Computers in Education.

Vatrapu, R., Reimann, P., **Johnson, M.** & Bull, S. (2013). An Eye-Tracking Study of Notational, Informational, and Emotional Aspects of Learning Analytics Representations, Proceedings of Learning Analytics and Knowledge.

Bull, S., Wasson, B., Kickmeier-Rust, M., **Johnson, M.D.**, Moe, E., Hansen, C., Meissl-Egghart, G. & Hammermuller, K. (2012). Assessing English as a Second Language: From Classroom Data to a Competence-Based Open Learner Model, International Conference on Computers in Education.

Johnson, M. & Bull, S. (2011). Optional Finer Granularity in an Open Learner Model, in G. Biswas, S. Bull, J. Kay & A. Mitrovic (eds), Artificial Intelligence in Education, Springer-Verlag, Berlin Heidelberg, 478-480.

Reimann, P., Bull, S., Halb, W. & **Johnson, M.** (2011). Design of a Computer-Assisted Assessment System for Classroom Formative Assessment, CAF11, IEEE.

Bull, S., Gakhal, I., Grundy, D., **Johnson, M.**, Mabbott, A. & Xu, J. (2010). Preferences in Multiple View Open Learner Models, in M. Wolpers, P.A. Kirschner, M. Scheffel, S. Lindstaedt & V. Dimitrova (eds), Sustaining TEL: From Innovation to Learning and Practice, EC-TEL 2010, Springer-Verlag, Berlin Heidelberg, 476-481.

Bull, S., Mabbott, A., Johan, R., **Johnson, M.**, Lee-Shim, K. & Lloyd, T. (2010). Interfaces for Inspectable Learner Models, in Aleven, V., Kay, J. & Mostow, J. (eds), Intelligent Tutoring Systems 2010, pp.275-277.

Johnson, M. & Bull, S. (2009). Belief Exploration in a Multiple-Media Open Learner Model for Basic Harmony, in V. Dimitrova, R. Mizoguchi, B. du Boulay & A. Graesser (eds), Artificial Intelligence in Education 2009, IOS Press, Amsterdam, 299-306.

Bull, S., Ahmad, N., **Johnson, M.**, Johan, R., Mabbott, A. & Kerly, A. (2008). Adaptive Navigation Support, Learner Control and Open Learner Models, in W. Neijdl, J. Kay, P. Pu & E. Herder (eds), Adaptive Hypermedia and Adaptive Web-Based Systems, Springer-Verlag, Berlin Heidelberg, 275-278.

Workshops and Interactive Events

Demmans Epp, C., Bull, S., & **Johnson, M.D**. (2014). Visualising Uncertainty for Open Learner Model Users, CEUR Proceedings associated with UMAP 2014.

Bull, S., Kickmeier-Rust, M., Meissl-Egghart, G., **Johnson, M.D.**, Wasson, B., Alotaibi, M. & Hansen, C. (2013). Interactive Event: From a Virtual Village to an Open Learner Model with Next-TELL, AIED 2013.

Johnson, M.D., Bull, S., Kickmeier-Rust, M. (2013). Student Competency Visualisation for Teachers, 3rd International Workshop on Teaching Analytics, EC-TEL 2013.

Bull, S., Wasson, B., **Johnson, M.D.**, Petters, D. & Hansen, C. (2012). Helping Teachers Effectively Support Group Learning, in J. Kim & R. Kumar (eds), Workshop on Intelligent Support for Learning in Groups, ITS 2012.

Reimann, P., Bull, S., Hesse, F., Cerniak, G., Luckin, R., Vatrapu, R., **Johnson, M.D.**, Halb, W., Utz, W., and Kossowski, M. (2012) Symposium on Supporting Teachers in Capturing and Analyzing Learning Data in the Technology-Rich Classroom, ICLS 2012.

Johnson, M.D., Reimann, P., Bull, S. & Fujita, N. (2011). Epistemic Beliefs and Open Learner Models, Workshop on Personalization Approaches in Learning Environments, User Modelling, Adaptation and Personalization 2011.

Chapter 1

INTRODUCTION

This thesis investigates *"how is an open learner model with a drill down approach used and what is its impact on self-assessment and metacognition?"* It aims to establish some of the context around the application of a drill down approach in an open learner model (OLM) with specific reference to deployment in a self-directed educational setting, and in particular considers the impact of the technology on the accuracy of student self-assessment and support for tasks that are regulatory to cognition.

Metacognition is 'thinking on thinking' or 'cognition on cognition'. Literature suggests that metacognition affects learning and that good learners reflect on how to learn, that self-assessment is a regulatory practice which contributes to metacognitive development, and that metacognition includes the regulation of cognition, which comprises tasks such as comprehension and problem solving. This thesis considers student engagement with the OLM with particular reference to their use of a drill down approach, which is used to structure and externalise the learner model to the learner for these purposes.

Several assumptions are made about the question, which are explored in the literature review:

- Aspects of metacognition are measurable.
- Mechanisms exist for measuring impact of self-assessment.
- A drill down approach is suitable for an open learner model.
- Self-assessment and metacognition are areas where an OLM should have impact.

This thesis then proceeds to propose and implement the design for technology to explore the use of a drill down approach in an OLM, and then responds to the research question in three parts, asking:

- How is an open learner model with a drill down approach used?
- What is the impact of an open learner model with a drill down approach on self-assessment accuracy?
- What is the impact of an open learner model with a drill down approach on the support for the regulation of cognition?

This thesis contributes:

- An open learner model with a drill down approach is used consistently across student use with a baseline pattern of activity of short bursts of questions followed by immediate inspection of the area of the OLM that is updated. Drill down is used conditionally to allow students to complete detailed inspection of problems that are apparent. The drill down structure also allows students to gain an overview of their learning without requiring all the detail of the model.
- The impact of an OLM with a drill down approach on self-assessment accuracy is that while there is no universal increase in self-assessment accuracy, some participants do improve in self-assessment accuracy during use and these students make greater use of drill down. These learners can also be accurately identified by asking students to give an indication of their own self-assessment abilities.
- The impact of an OLM with a drill down approach on the support for the regulation of cognition is that it can readily facilitate and evidence student actions that are consistent with problem solving, as a subprocess that facilitates the control of learning (metacognition). Problem solving behaviour is part of the baseline interaction of students interacting with an OLM of this type and drill down is used in a way as to support the conditional access of information that is required specifically when the detailed inspection of problems is to be undertaken.

To respond to the research theme, this thesis is organised as follows:

Chapter 2 clarifies the initial assumption in the literature that aspects of metacognition are measurable, and it establishes the role of self-assessment and self-directed learning as part of the learning process. Chapter 3 then considers open learner models, information granularity and a drill down approach, and the support that that these have for metacognition. Chapter 4 builds on both these components to look at applied examples of how a drill down approach has previously been realised in open learner modelling and design aspects related to this.

In Chapter 5 the thesis introduces the approach to an evaluation to investigate the research theme and identifies the sub-questions and methods used throughout the analysis. Chapter 6 then introduces the design and implementation of technology specifically built as part of this research for students to use during the evaluation.

The evaluation itself is ordered in two chapters: Chapter 7 reports on a broad consideration of aspects of the research question with 4 experimental groups in 1 hour studies, and Chapter 8 reports on student use within a longitudinal study over an 8 week period.

Finally, Chapter 9 presents a summary and suggests future work.

Chapter 2

METACOGNITION, ASSESSMENT AND EDUCATION

This chapter considers the concept of metacognition as an important notion underlying education. Metacognition underpins much of the design of educational technology and indeed of open learner models (Chapter 3), aiding students in becoming effective learners in a self-directed learning context. Important aspects of this include the provision of formative- and self-assessment opportunities within learning.

2.1 Metacognition

Metacognition is often referred to as the concept of thinking about the way in which we think. Mahdavi (2014) outlines that while the term has many different facets, most simply metacognition can be summarised as a “critical analysis of thought”, “thinking on thinking” or “cognition on cognition” (Wellman, 1985; Anderson, 2008; Livingston, 1997). Flavell (1976) considers this as concerning knowledge of “one’s own cognitive processes, products or anything related to them”, as a higher-order process, something that acts on cognitive information or artefacts, which may be knowledge of a person, task, strategy or experience.

Veenman et al. (2006) collects a variety of different terms that have emerged from the field including metacognitive beliefs, metacognitive awareness, metacognitive experiences, metacognitive knowledge, feeling of knowing, judgement of learning, theory of mind, higher order skills, learning strategies and self-regulation; some of these are used in contradictory ways (Weinert, 1987). Schraw and Dennison (1994) summarise that metacognition can comprise:

knowledge about cognition, which can be categorised as:

- declarative: knowledge about self and about strategies.
- procedural: knowledge about how to use strategies**.**
- conditional: knowledge about when and why to use strategies.

regulation of cognition, including subprocesses that facilitate the control of learning, for example through key skills such as:

- planning.
- information management strategies.
- comprehension monitoring.
- problem solving/debugging strategies.
- evaluation.

Arguably, self-appraisal and self-management are the core aspects of metacognition (Paris, 1990). This includes both student awareness of their own knowledge, (including strengths, weaknesses and beliefs), and their ability of regulate their own actions in the application of that knowledge (Tanner and Jones, 1999). Those who might be considered the most effective learners would be self-regulators (Butler, 1995) and good self-assessors (Boud and Falchikov, 1989), with the ability to think critically in terms of what they have done and actions that they have taken (van Aalst et al., 2015). There is however evidence that students are not always good at self-evaluation (Boad and Falchikov, 1989; Mitrovic and Martin, 2007), that metacognitive learning is difficult (Zohar and Dori, 2003), that the approach is heavily dependent on student ability to accurately assess what is known (Schoenfeld, 1987) and that students need support in this (Mitrovic and Martin, 2007; Bull and Kay, 2013).

While some habitual behaviour and lower level consciousness might still be metacognitive (Veenman et al., 2006), enhancing metacognitive awareness in students is important (Bull and Kay 2013; Schraw and Dennison, 1994) and central to knowledge building (Scardamalia, 2002). It has important effects on cognitive tasks in terms of acquisition, application, comprehension, critical thinking and problem solving (Hartman, 1998). Student reflection in this dimension is important to foster the self-regulation of cognition (van Aalst et al., 2016), and it is not something that is innate, but is something that needs to be fostered (Veenman et al., 2004; Boud and Falchikov, 1989).

With reference to the regulation of cognition as a component of metacognition, tasks such as comprehension monitoring and problems solving have defined behavioural aspects. Namely, problem solving is aiming to achieve a goal without initially knowing the solution method (Mayer, 2013) and this will occur in a series of phases including:

- problem identification: recognising, identifying or defining the problem.
- structuring the problem: careful and inspection, observation, information gathering.
- looking for possible solutions: a range of possible actions or strategies.
- making a decision/solution implementation: carry out a course of action.
- monitoring/seeking feedback: evaluate success of action, further inspection.

These may be particularly identifiable in domains such as mathematics where there is often one solution to a problem (Blanchared-Fields, 2007) and cognitive processes for problem solving may include (Wang and Chiew, 2010): *hypothesis testing* (assuming an explanation and trying to prove/disprove it), *means end analysis* (choosing actions that move the learner closer to a goal), *morphological analysis* (observing outputs on an entire system), *root cause analysis* (identifying the cause of a problem) or *trial and error* (repeated attempts at a possible solution until a correct one is found). Overall, the idea of problem solving is a "redirection in thinking" (Smyth et al., 1994) and can be part of routine methods or habitual ways of approaching a problem. There are three defining properties to problem solving: (1) it is purposeful; (2) it involves cognitive rather than automatic processes; and (3) a problem only exists when someone lacks the relevant knowledge to produce an immediate solution (Eysenck and Keane, 2005).

- Metacognition comprises: knowledge about cognition (declarative, procedural, conditional); and the regulation of cognition.
- Student reflection and self-assessment are important to metacognitive activity, and tasks such as comprehension, critical thinking and problem solving are core to the regulation of cognition.
- Students are not always good at self-evaluation and metacognitive learning is difficult.
- Problem solving has phases with associated identifiable behaviours (e.g. including problem identification) and strategies (e.g. trial and error).

Learner reflection and self-assessment are identified as important contributors to metacognitive development, both of which this thesis now explores in greater detail.

2.2 Reflection and Self-Assessment

An important component of metacognition as a higher-order process (Flavell, 1975) is the capability of reflection: "conscious exploration of one's own experiences" (Boud et al., 1985), making thinking processes explicit (Silver, 2013). Reflection is a vital element in any form of learning, and one that can occur before, during and after an activity to which it may relate (Boud and Molloy, 2013). Reflection may be thought of as: reflection-in-action, whist undertaking activity (time critical); reflection-on-action, post hoc; or, in a more meta form, reflection-on-reflection itself (Schon and DeSanctis, 1991). Reflection is also core to part of the learning cycle (Kolb, 1976).

While reflection is common, it is false to assume that it always occurs effectively (Boud and Molloy, 2013). However, learners may realistically improve their effectiveness by paying attention to their reflection (Li and Kay, 2005), and students should be given opportunities to think in terms of strategy, and to reflect on their learning (Black and Wiliam, 2006).

In terms of metacognition, reflection can take different forms such as, critiquing, planning, questioning, justifying, and explaining (Dewey 1933; Boud and Molloy 2013), all of which are vital activities for active participation in learning. Boud and Molloy (2013) describe reflection as an active process of discovery, and one that can often lead to unexpected outcomes. Indeed, fundamental to many aspects of reflection is the concept of assessment, making a judgement about a specific state, potentially making "formerly unconscious, intangible or reflexive processes or events explicit" (Desautel, 2009). In terms of metacognition, this is then an analysis of the discrepancy between one's mental model and an observation. This is, more specifically, self-assessment: a specific subset of reflection (Boud and Molloy, 2013). Self-assessment is essential to the concept of autonomous learning (Cassidy, 2006), is a regulatory practice that contributes to metacognitive development (Desautel, 2009), encourages reflection by promoting the student to the role of their assessor, and should be supported (Mitrovic and Martin, 2007). Self-assessment is closely linked to self-monitoring, through which a person compares products with standards to determine if objectives have been met or if further work remains to be done (cognitive evaluations) (Azevedo et al., 2010).

- Reflection (core to metacognition) can have different points of initiation, with different temporality, and cognitive forms (e.g. critiquing, planning, questioning, justifying, and explaining). It is a process of discovery for the learner.
- Reflection may not always occur effectively, but students may improve in effectiveness if given opportunities (interventions) through which they may perform reflective thinking.
- Self-assessment is a sub-set of reflection, contributes to metacognitive development, and is important in self-directed learning. It fits well with formative assessment opportunities.

Self-assessment is a regulatory practice that contributes metacognitive development and formative assessment opportunities may provide opportunity for reflection and metacognition to occur. In the following section this thesis looks at formative assessment in more detail.

2.3 Formative Assessment

Assessment types fall into different categories. Section 2.2 has considered the idea of self-assessment as a reflective practice in which learners are able to make assessments about themselves. Similarly, assessment of students may be considered as summative or formative. Assessment is summative if it is to inform on student progress, described by Crooks (2011) as "a summary of students' descriptive content knowledge", or formative if it is "assessment with the intention to help learners improve content knowledge and/or skills" (Price, 2015), although there is some ambiguity in terms of a consistent formal definition. Formative assessment aims to go beyond establishing the state of student knowledge and to generate feedback with the intention to augment or expedite learning. It is formative assessment that is of particular interest to this thesis from the perspective of metacognitive practices, as it provides a point of initiation for metacognitive activity, and something tangible that can be used for furthering the student's learning and enhancing their understanding.

Formative assessment has been described as a continuous and systematic process to gather evidence and provide feedback during learning (Heritage et al., 2009), as something that can often be immediately used (Shepard, 2008), and as something that allows the appreciation of standard and qualities of work, with a view to producing improved work (Boud and Molloy, 2013). The ultimate aim for formative assessment is to help identify where the student is within their learning and then identify gaps and what he or she needs to do next to further his or her knowledge or skills (Pinchok and Brandt, 2009). Initiatives such as Assessment *for* Learning (see Black and Wiliam, 1998) underline that assessment is part of the learning process. The benefits of student involvement in assessment, and also the leverage it has over student self-esteem and motivation, are iterated by Black and Wiliam (1998). It has also been argued that formative assessment opportunities themselves need to be planned (Popham, 2008). Calls have been made to take this beyond the concept of feedback into feedforward, specifically an artefact (physical or verbal) that looks forward, with content about how the student might improve: Whitelock (2010) describes this as "advice for action", and further states that feedback is not just for after a task, but something that provides hints before a task is commenced.

Hattie and Timperley (2007), summarised in Ras et al., (2016), suggest that formative assessment feedback occurs on four levels:

- Task level – feedback related to whether the task/activity is correct or not.
- Process level – creation of a product or the accomplishment of a task.

- Self-regulation – judgement of self-regulation of confidence skills with regard to the task (see also Section 2.4).
- Self-level – elements that are personal and potentially unrelated, such as praise.

Likewise, Chappuis (2009) proposes formative assessment for learning strategies including:

- Provision of a clear understandable vision / learning target.
- Examples and models of strong and weak work.
- Regular descriptive feedback.

and strategies which include skills for self-directed learning (see also Section 2.4):

- Teach students to self-assess and set goals.
- Teach students to focus on one target or quality at a time.
- Teach students focussed revision.
- Engage students in self-reflection, and to keep track of and share their learning.

Black and Wiliam (1998) in their comprehensive review advocate the benefits of formative assessment, and state the consistent positive effects and improvements resulting from its use, even if there are sometimes great difficulties putting theory into practice and translating it to the educational setting (Black et al. 2005). Overall, how effective this might be relies on students' (metacognitive) skills to perceive the gap between where they are and where they need to be (Biggs, 1998; Nicol and Macfarline-Dick, 2006).

With consideration of metacognition, (formative) assessment is an initiation point for an episode of reflection and metacognitive activity and requires a certain amount of self-assessment (and hence reflection) to be able to interpret it and make judgements (Nichol, 2009). Desautel (2009) cites ways of encouraging successful student self-assessment, including student-generated rubrics, portfolios, contracts, and goal setting, and the practice of self-assessment is something that is important for further involving students in their own learning (Jackson and Larkin, 2002; Yancey, 1998) and in the regulation of cognition.

Technology that facilitates formative assessment also promotes the aims of metacognition, reflection and self-assessment in a way that uses data driven decision making to enhance and transform assessment practice (Redecker and Johannessen, 2013) and opens doors for new practices and feedback types (Ras et al., 2015). “It is the learners and teachers as human actors who ultimately determine the formative effects of engaging with technologies, but technologies can shape the potential for this to happen” (Pachler et al., 2010).

- Formative assessment (as a learning-based intervention method to facilitate both the reflective and regulatory aspects of metacognition) is systematic/continuous feedback/feedforward aiming to improve an aspect of student learning. This is part of the learning process and should be planned.
- Effectiveness of formative assessment relies on students’ metacognitive skills.
- Formative assessment (as an intervention) may be used to develop strategies for effective self-directed learning and regulation of self-assessment skills and might be enhanced by technology.

Formative assessment may be used as an intervention to promote and influence metacognitive activity and may be used to facilitate self-assessment practices. The effectiveness of formative assessment of this type may rely on students’ metacognitive abilities, particularly in terms of self-reflection and this is particularly applicable to self-regulation. As such, this thesis now considers formative assessment, in the context of self-directed learning.

2.4 Self-Directed Learning

Pintrich (2000) describes self-directed learning as an active process in which students attempt to monitor, regulate and control aspects such as their cognition, behaviour or motivation, underpinned by a strong focus on goals and context. This is arguably of key importance in fostering metacognitive skills. Gabrielle (2003) also states that self-direction is part of a student's desire to learn beyond the requirements of their course. Furthermore, this is something that is core to lifelong learning (Boud, 2000), and something that fits very well with the empowering nature of formative assessment feedback that allows learners to "plot, plan, adjust, rethink and exercise self-regulation" (Hattie and Yates, 2014), and to allow learners to make decisions based on evidence (Black and Wiliam, 2009; Stiggins et al., 2006). Self-directed learning stems from a firm underpinning that students should identify their goals, strategies and models and be able to evaluate their decisions and outcomes (Knowles, 1975, in Mello, 2016). This is, in essence, formative assessment of one's own learning practice, and much research has focused on how students are able to regulate their cognition, metacognition, motivation and engagement with tasks (Azevedo et al., 2010), and in doing so self-directed learning encompasses aspects of critical thinking and has the potential to encourage deep learning (Handelsman et al., 2004, in Mello, 2016).

Grow (1991) cites four stages of self-directed learning, in the context that learner autonomy of this type is something that needs to be fostered in students. Learners may be:

- dependent (e.g. receives coaching with immediate feedback; overcomes deficiencies).
- interested (e.g. goal-setting and has a learning strategy).
- involved (e.g. discussion with another, equal participation).
- self-directed (e.g. individual working, self-direction towards attainment of a goal).

It is also acknowledged that not all students may be at the same level of self-regulation, but those who are the most effective undertake self-directed learning (Butler and Winne, 1995). Grow (1991) postulates that those who are closer to the dependent/interested stages are likely to need an authority, an expert or an entity to keep them motivated, whereas those who tend more towards the self-directed/involved stages may instead need someone to facilitate their learning, or to delegate the responsibility to them. As such, projects requiring a degree of independence are ideally suited to self-directed learners, and projects that allow immediate correction of problems are suited to dependent learners, although there is a degree of cross over between these two aspects. Gabrielle (2003) also extends this to consider that learners may be at different levels at different times and with different topics; it is heavily contextual.

Effective self-directed learning can occur in four phases: task definition, goal-setting and planning, studying tactics, and metacognitive adaptations (Winne and Hadwin, 2008 in Azevedo et al., 2010). Thus, it takes the process from the inception of a task through to a reflective analysis of whether it has been achieved and how improvement might be made. There is suggestion though that self-regulation within learning is not always effective: it might be affected by error-carried-forward behaviour (e.g. if a learner believes that a task is too easy), and might be affected by cognitive load, inefficient use of cognitive strategy, lack of metacognitive control, or lack of prior knowledge (summarised in Azevedo et al., 2010). Shute (2008) also emphasises the point that feedback of any type should be in manageable chunks, to avoid cognitive overload, and should be actionable.

Self-direction relies greatly on self-assessment and formative assessment. Ras et al. (2015) emphasise that in the context of self-directed learning students need to be able to comprehend the gap between their current learning state and their target goal, and have the freedom to make decisions in terms of their learning, and to also take responsibility for their learning. Citing Good (2011), Ras et al. (2016) state that strategies interact tightly with learning context and content. Shute (2008) highlights several relevant points in this context, in terms of effective strategy for feedback, namely, that:

- delayed feedback is more effective for simple tasks and immediate feedback for complex tasks.
- facilitative feedback is more useful for high achieving learners, when delayed.
- directive (corrective) feedback should be delivered immediately to low-achieving learners.
- low performers profit from scaffolding and using feedback where the correct answer is given, or an elaboration of the correct answer is provided.
- for high performing learners, verification feedback might be sufficient.
- low learning orientation leaners need specific goal-directed feedback.

This thesis considers that suitable formative assessment feedback may also be provided by online resources (Mello, 2016; Markauskaite and Jacobson 2015; Hattie 2008; Sharpe et al., 2006; Chudowsky and Pellegrino, 2003; Glaser et al., 2001), such as providing individualised feedback for reflection on knowledge in and out of the classroom (Glaser et al., 2001), to enrich assessment situations and support problem solving skills (Chudowsky and Pellegrino, 2003) and to provide opportunities for students' independent and self-directed learning, where changes in course structures leave students requiring additional support (Mello, 2016). It is acknowledged that interaction behaviours with online formative assessment technologies can range from approaches that are conceptually astute, through to incomplete and unsystematic behaviours (Thompson and Reimann, 2010, in Markauskaite and Jacobson 2015): not all technology engagement strategies lead to consistently successful outcomes, and so there is an important requirement the nature of learners' interactions with technology are better understood. That students can conduct authentic enquiry, such as in the self-directed learning context, is an overarching goal of the science, technology, engineering and mathematics (STEM) curriculum and a critical skill in the 21st century (Pellegrino et al., 2014, in Markauskaite and Jacobson, 2015). Hattie (2008) states that the strongest effects of technology were reported when computers supplemented traditional teaching, when students were assumed to have control over their learning situation (pacing and mastering new material) and when computers were used to provide adaptive feedback.

This thesis therefore builds on the foundation that self-directed learning can be facilitated by educational technology with a strong underpinning in terms of the promotion of metacognitive development, and that there is a desire to understand more about how learners make use of the technology.

- Self-directed learning is an active process involving the regulation of cognition, and aspects of student self-assessment, which are the two key aspects of metacognition. The most effective learners undertake self-direction, but not all are able to do this.
- Strategies exist for the best point for learners at different levels to receive feedback (intervention) and technology has the potential to deliver this.
- There is a strong focus on goals and context. Intervention methods (e.g. formative assessment) are a decision point.
- Issues such as cognitive load, information volume, lack of metacognitive control, and lack of prerequisite knowledge will have an effect on student ability to undertake self-direction.

Effective self-direction encompasses many aspects of metacognition, using feedback as an intervention and a decision point within learning. The suggestion from the literature is that there should be an impact on metacognition and self-assessment through learning of this type, but that it should not automatically be assumed to be taking place, nor should it be expected to be a uniform effect.

2.5 Summary

This chapter has highlighted that some aspects of metacognition are measurable, that metacognition can be impacted by learning, and that there are many different dimensions to metacognition, including those relating to learner reflection and those relating to the regulation of cognition. It has established that self-assessment is a regulatory practice that contributes to metacognitive development and also that the use of a self-directed learning context has potential to lead to metacognitive effects. The chapter has also highlighted a series of student tasks that can facilitate the regulation of cognition, including comprehension, planning, critical thinking or problem solving, and a suggestion that technology-based solutions may facilitate these.

Furthermore, this chapter has indicated that while metacognitive effects should be observed, there is no certainty that they should occur, and that there are great differences between students, so it is valid to continue to confirm situations in which metacognitive benefits might be achieved, and situations in which tasks and processes that contribute to this are effective. It has highlighted that the design of the technology (as a formative assessment-based intervention) is important, not least in terms of issues such as cognitive load, volumes of information, timeliness, and the context of the origin of the information. Similarly, it is important to underline that the content of the formative assessment is related to some aspect of the learner, such that reflection might occur.

This chapter therefore contributes the following requirements to the design of its evaluation, which regard tasks relating to the regulation of cognition, measurement of self-assessment, and the use of self-directed learning:

Table 1: requirements.

ID	**Description/Rationale**	**Supporting Sections**
R1	Tasks such as evaluation, comprehension monitoring and problem solving/debugging relate to the regulation of cognition, as a subset of metacognition that is more measurable. The technology should allow for evaluation of aspects of these.	Metacognition
R2	Self-assessment is a regulatory practice that contributes to metacognitive development. Consideration should be given for how this changes across time to understand its impact.	Self-Assessment Metacognition
R3	An appropriate point for OLM use is during attempts at self-directed learning. This is a point during which impact on self-assessment or aspects of metacognition might be expected in technology such as this.	Self-Directed Learning Formative Assessment Metacognition

This thesis next moves forward in Chapter 3 to consider the design of open learner models as a method through which learners may access information about themselves and their learning.

Chapter 3

OPEN LEARNER MODELLING, DRILL DOWN APPROACHES AND INFORMATION GRANULARITY

The literature of Chapter 2 highlights that metacognition should affect learning, and that formative assessment in a self-directed learning setting is suitable for promoting metacognitive activity. Tasks such as problem solving can be measurable aspects of the regulation of cognition, and self-assessment is important as a measurable regulatory practice that contributes to metacognitive development. Issues such as information volume, detail, content and temporality are also important when these practices are facilitated through technology.

Building on these requirements, the thesis considers two technology-related aspects of the research question, that of:

(i) open learner models as technology that supports self-assessment practice, promotes the reflective aspect of metacognition and allows for the regulation of cognition though learner access to formative assessment information about themselves and their learning, and

(ii) the use of a drill down approach as a structural solution to the presentation of information, to allow some insight into the types of tasks that students may be attempting, in the context of self-direction.

This chapter considers both areas of literature and then this is extended in Chapter 4 to look specifically at the existing use of drill down approaches in open learner modelling, with view to implementing requirements for the design of a suitable investigatory method (Chapter 6). This chapter begins by summarising what learner models are and their origins in intelligent tutoring systems.

3.1 Intelligent Tutoring Systems

Intelligent tutoring systems (ITS) are pieces of adaptive educational technology that provide personalised and specific tutoring customised to the needs of the learner (Baker, 2016). Born out of earlier work in computer assisted learning (CAL), their design is built around a tripartite architecture containing each of the following principal elements (Self, 1999):

(i) Domain model. This contains the information about what is to be taught. It is a reference corpus containing a range of expertise, concepts, rules, relationships, knowledge or methods for problem solving. This is sometimes referred to as the 'expert model'.

(ii) Learner model. This is the system's understanding of the learner who is to be tutored. It is built during interaction, with inferences about the learner, in the context of the capabilities and expert information held in the domain model.

(iii) Teaching strategies. The tutorial strategy element draws on information from both the learner model and the domain model to regulate how the student is instructed, adapts interaction according to the needs of their learning, and decides what information to acquire from the learner. The strategies it contains act dynamically in the context of student action.

Further to this, the user interface (UI) is sometimes considered as the fourth element. The UI is important to the user-system relationship, as it provides the means of communication. In the traditional ITS approach the learner model is available only to the system itself, and not to the learner, and the system very much retains control. The learner modelling element has been prime in the continuation of this research, and work of the last few decades has considered making this available (open) to the learner, for a range of metacognitive benefits, formative assessment opportunities and as support for self-directed learning (see Bull and Kay, 2010 for an overview). Work reported in this thesis builds on this research strand.

- Intelligent tutoring systems contain a learner model component, well suited to collating information about student learning, and can express this in terms of the domain.
- The user interface is important to the user-system relationship and should be carefully designed depending on the aims of the system. The user does not automatically have access to everything.

The learner model component of an intelligent tutoring system can very easily express information about learners, which is an important requirement highlighted in Chapter 2. Next, the thesis considers what form learner models might take.

3.2 Learner Models

Learner models reflect the learner's current learning needs and are usually dynamically updated throughout periods of interaction with a system (Bull and Kay, 2016). The contents of the model may include a diverse range of elements such as competencies, knowledge, skills, abilities, aptitudes or 21st Century skills (Reimann et al., 2011). In contexts where different aspects of the learner may be required to be considered it might also include such elements as affective states, collaborative interactions, engagements with activities, or student-, teacher-, or peer- assessments (Van Labeke et al, 2007; Johnson et al., 2013b). The learner model's purpose here is to be able to provide answers to questions about the learner's current needs and is at all times a predictive diagnosis and estimation of the student within these dimensions. It is important to note that the knowledge state of the learner can both increase (learning) and decrease (forgetting) (Kump et al., 2012).

To fulfil its purpose, the learner model can be constructed using a variety of methods from simple weighted numerical algorithms (e.g. Zapata-Rivera and Greer, 2004), through to machine learning techniques that can learn to solve problems within the context of the domain (e.g. Sison and Shimura, 1998), and constraint based procedures that verify actions through the non-violation of constraints (e.g. Mitrovic and Ohlsson, 2016). (See also Woolf, 2008 for an overview of learner modelling methods.) Learner models may be updated upon receipt of new information or indeed may follow an active learner modelling approach (McCalla, 2004) whereby computation is completed at the point of access. This variety of techniques also implies that the underlying representations of the model also vary, each suited to the model's purposes and aims. These representations might be numerical, symbolic, mathematical, logic-based, production rules or networks, for example (Bull and Kay, 2016). This in many cases may mean that the full volume of information underlying the model is of a much greater richness than that which is required within its educational role, and also may implicitly incorporate some aspects of domain information (e.g. structural relationships). The granularity of the information is also an important consideration, particularly for aspects of uncertainty and pedagogical decision making (McCalla and Greer, 1994).

In an increasingly digitised world, sources of information for the learner model have the potential to be numerous, and models built from multiple sources are becoming more commonplace (Reimann et al., 2011; Rayborne and Regan, 2011). They may take into account a fuller range of activities, data collected as students interact, online activities (e.g. sent through an API), traditional feedback, or student/peer assessments. This also highlights the practical consideration that learner models may sometimes have to persist with missing, conflicting or incomplete data, which may not always be apparent in surface-level aggregations of items within a model. This is core to Self's (1990) call for the learner model to be opened to the learner "*in order to provoke the student to reflect upon its contents and to remove all pretence that the ITS has a perfect understanding of the student*".

- Learner models provide information for (the system) responding to the learner's current needs, by making a prediction, based on a potentially rich body of information.
- Information can exist on different levels of granularity/abstraction, accessed at a suitable level as to answer a specific need. This can lead to conflicting and incomplete information which is not always apparent at surface level.

This section highlights that a learner model can readily contain information required to support reflective activity (as described as critical to metacognition in Chapter 2). It also underlines that not all issues are apparent at surface level representations, that learner models can hold large volumes of information, and that they are traditionally intended for use by the system. This thesis now considers the case for opening the learner model to the learner (making it inspectable) in terms of benefits for metacognition, formative assessment and self-directed learning (Section 3.3), and what form it might take (Section 3.5).

3.3 Open Learner Models

An open learner model is a model of the learner that is able to be accessed by the learner (or other educational stakeholder) in an interpretable form (Demmans Epp and Bull, 2015; Bull 2016). This can also be referred to as an open student model, and the open representation of the learner model can be termed a view or visualisation (Mathews et al., 2012). Some models are opened for inspection only (whether the visual methods are interactive or more static) (e.g. Johnson et al., 2013b), while others embody methods and mechanisms for interactive maintenance (e.g. Ginon et al., 2016; Mabbott and Bull, 2006; Bull and Pain, 1995), whereby the learner has some level of control over the contents of the model itself.

The main reasons and stated purpose for opening the learner model are based around attempts to facilitate metacognitive benefits. This makes this approach highly relevant to the research question. Building on the ideas and concepts covered in Chapter 2, such motivations to open the learner model to the learner are to promote learner reflection, provide formative assessment opportunities, encourage self-assessment, support planning, monitoring and learner autonomy within learning (self-directed learning), facilitate collaboration/ competition, support navigation, and improve the accuracy of the learner model (as summarised by Bull and Kay, 2013). Furthermore, this addresses each learner's right to see information held electronically about them and increase learner trust in the information (Ahmad and Bull, 2008).

The open learner model is able to provide individualised feedback to students and the data visualised relates directly to their understanding, skills and competencies (Morales et al, 2008), and any other aspects upon which the model is able to provide information, such as affective states (e.g. Van Labeke et al., 2007).

Learner models have been opened to different age ranges of students, including learners at university level (e.g. Mitrovic and Martin, 2007; Bull and Mabbott, 2004). Some instances of OLMs are used as learning resources in their own right, being made available independently from any wider educational environment, also potentially embodying an element of domain independence, or as a collection node for data from multiple sources (Reimann et al., 2011). Taking this approach potentially allows for students to assume more responsibility for, and control over, learning decisions and the development of metacognitive skills (Bull et al., 2014), whether this may originate from a system that has tasks or questions (Bull and Britland, 2007) or from other educational systems such as e-portfolios (Raybourn, 2011). OLMs have been deployed in a variety of domains from computer programming (e.g. Mabbott and Bull, 2006), and mathematics (Long and Aleven, 2013, Van Labeke et al., 2007) through to language learning (Dimitrova, 2003) and music (Johnson and Bull, 2009). The majority are in STEM domains, which very naturally lend themselves to the modelling process, although there is increasing interest in work with more ill-defined domains (Mitrovic and Weerasinghe, 2009).

A variety of research in recent years has been completed on open learner models including: learner model negotiation and interactive maintenance (Kerly et al., 2008; Ginon et al, 2016); issues of trust (Ahmad and Bull, 2008); OLMs facilitating collaborative behaviour (Alotaibi and Bull, 2012; Brusilovsky et al., 2011); issues of accuracy (Long and Aleven, 2013), support for self-assessment accuracy (Mitrovic and Martin, 2007); and the opening of the learner model in new and novel visualisation types (Johnson and Bull, 2009; Lloyd and Bull, 2006). What remains at the heart of this research is how each dimension can support students in the most effective ways, particularly in terms metacognitive aspects (Bull and Kay, 2013) and formative- and self-assessment (Kerly et al., 2008). Research suggests that there are significant learning benefits from interactions with open learner models, especially for weaker students (Mitrovic and Martin, 2007), although these learners may often struggle to interpret complex aspects of the learner model and potentially need the most support (Brusilovsky et al., 2011). Arguably the initiation points for metacognitive activity and reflection can be many within the technology and can vary from learner to learner, in alignment with individual differences, learning preferences or the extent to which metacognitive skills are already developed by the learner. The complexity and size of the learner model are also of key consideration, where these factors might influence student interaction and engagement, which from the self-directed learning perspective is essentially part of a sense-making exercise.

- Learner models are opened (made accessible) to the learner with the intention of facilitating self-assessment practice and metacognitive benefits, providing individualised feedback to students: these are open learner models (OLMs). They also make explicit information that is prerequisite to the completion of tasks that are metacognitive in nature.
- Significant learning benefits can be seen, but OLMs are not useful to all students all the time: for example, some often struggle to interpret complex aspects of the model.
- How open learner models best support students in terms of formative assessment, metacognitive practice and self-directed learning are ongoing lines of enquiry.

Facilitating metacognition is an important line of research in open learner modelling, and reasons for opening the learner model to the learner are centred on potential metacognitive benefits. In a self-directed learning context, OLMs can act as an intervention in the learning process through providing a formative assessment opportunity. They provide a body of information about the student that may be used for reflective activity, informing tasks appropriate to the regulation of cognition and for self-assessment. Literature also highlights the diverse content they OLMs typically hold, and different levels of abstraction to which they may be opened. This theme must be carefully considered, as the literature of Chapter 2 suggests interventions such as formative assessment are decision points for students, that issues such as cognitive load will affect this, and that for reflective activities (e.g. planning, justifying, explaining) students require sufficient and accessible information to be available to undertake this.

It is worthy of note that many aspects of learning analytics also take a similar approach to the above, providing students with information about their learning. This thesis next briefly considers the cross over with the field of learning analytics.

3.4 Open Learner Models and Learning Analytics

Learning analytics is a relatively new field within the educational domain, which synthesises aspects of related areas such as academic analytics, action research, educational data mining, recommender systems and personalised adaptive learning (Chatti et al., 2012). Considerable interest has emerged in the visualisation of educational data (Tervakari et al 2014; Verbert et al. 2014). While definitions can be varied, it centres on the “use of intelligent data, learner-produced data and analysis models to discover information and social connections, and predict and advise on learning” (Siemens, 2010, in Chatti et al., 2012). In its core elements, this kind of information can be used effectively to make educational recommendations (Duval, 2011), and relates to tracking user activity and behavioural interaction, grown from Big Data, Educational Data Mining, and information visualisation techniques (Bull et al., 2016). In a similar approach to OLMs that externalise data from learner models, learning analytics display such data and information to users (e.g. Verbert et al., 2013), who may also be of different stakeholder types. It has been previously argued that open learner models can very naturally make this information available to different members of the educational setup, and that these can be considered as a specific application of visual learning analytics (Bull et al., 2014). It is now the case that learning analytic approaches are being discussed together with open learner models, due to their great commonalities and literature suggesting that they may both fruitfully benefit from each other (e.g. Durall & Gros 2014; Ferguson 2012; Kalz 2014; Kay and Bull 2015).

- Learning analytics and OLMs overlap in terms of their goals and aims.
- Learning analytics can provide a much more diverse level of information about students in terms of their general activity. Learner models are often more focussed on student knowledge, skills and aptitudes.

Both OLMs and learning analytics provide students with suitable information about their learning to potentially support them, and provide the opportunity for episodes of metacognitive activity, such as on strategies, attainment of goals or supporting planning activities. Arguably OLMs are a specific form of learning analytic. This thesis focuses more specifically on cognition with reference to the student's cognitive state, so both the system and student can assess the same variables. This is to support potential tasks relating to the regulation of cognition such as debugging and comprehension, and accuracy of self-assessment.

Having presented the concept of an open learner model, this thesis now considers ways in which the learner model may be opened to the learner.

3.5 Visual Representations of Open Learner Models

Learner models may be opened to the learner in a very diverse range of formats and underpinning this is the idea that the learners themselves have to use the visualised data to identify what they need to do to improve their learning. This is part of the learning process (Bull et al., 2014), is an effective way of supporting sense-making, particularly where the visual methods are able to synthesize complex information for learners to quickly understand (Heer and Agrawala, 2008), and also can bring together data from multiple sources (Bull et al., 2013). The underlying learner model may often be considerably more complex than is able to be understood by its viewer, foremost as it is initially intended for use by the system itself.

The learner model may be opened in a diverse range of ways (Figure 1). These may include simple presentations (e.g. skill meters - Long and Aleven, 2013) or those that are more complex (e.g. Bayesian networks or 3D network structures - Zapata-Rivera and Greer, 2004); they may be textual (e.g. Bull and Pain, 1995) or graphical (e.g. Mathews et al., 2012); they may display structural relationships (e.g. network map Zapata-Rivera and Greer, 2007); or they may be quantised (e.g. Bull et al., 2013). Some visual methods may be interactive (e.g. treemap Brusilovsky et al., 2011), metaphorical (e.g. smilies - Kerly et al., 2008), display information across time (e.g. Ginon et al., 2016), be combined with other aspects of learning analytics (e.g. Johnson et al., 2015), be domain dependent (e.g. Johnson and Bull., 2009; Johan and Bull, 2009), or may open the learner model from different perspectives simultaneously (e.g. heatmap - Ginon et al., 2016). Over time, designers are becoming increasingly more creative in terms of OLM designs, although consistency in the way in which they're interpreted is an open issue (Mathews et al., 2012). Mabbott, 2009 also provides a more in-depth discussion of visualisation.

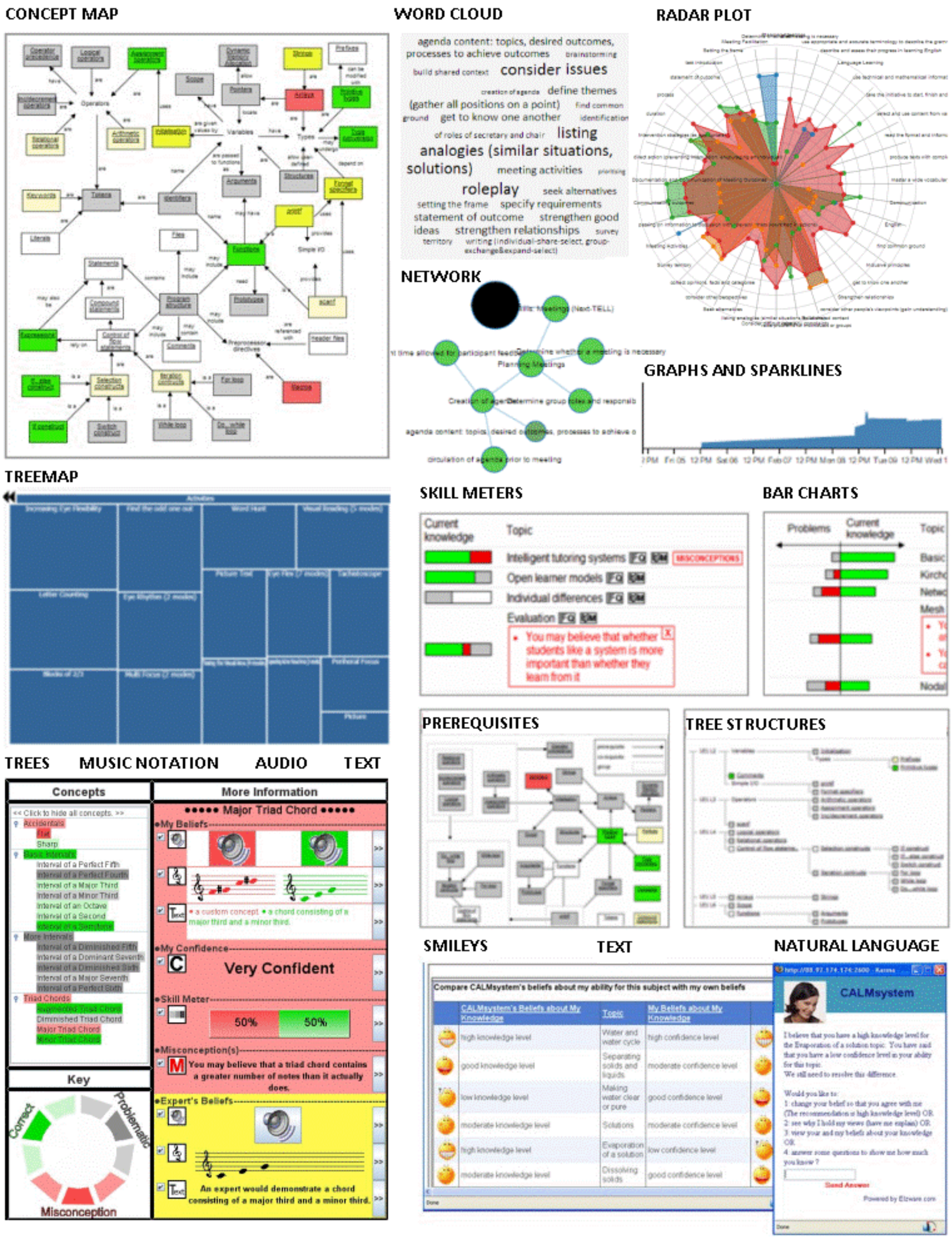


Figure 1: example OLM visualisations. Images reproduced with permission.

In all cases, the visualisations used within a system are selected for their suitability for purpose and for the ability level of the learners who will use them. For example, children need a simpler format, and so metaphors such as smiley faces may be appropriate (Kerly et al., 2008). Where the aims of the interaction are also to understand the structural elements of the model and how concepts interrelate, network diagrams may be more appropriate (Ahmad and Bull, 2008). Where the learner model is domain specific, a natural representation of the domain content may be appropriate (Johnson and Bull, 2009). It is also important to take into account the different preferences, prerequisite skills, backgrounds and aptitudes of a group of learners, and for these reasons presenting alternatives may be a useful approach (Mabbott, 2009). Where the same underlying content is used to drive different visual types, these may be considered isomorphic (Vatrapu et al., 2013): that is they present the same information in different ways.

Arguably the inspection of the learner model may be considered as something that can be task driven, and there will be different information requirements at different points during interaction. Those that require a quick answer at a high level may be more suited to a simpler model representation at a higher level of abstraction, whereas a specific point of enquiry may require a level of detail that may even be beyond the scope of that provided by the visualisation set. Issues of scalability, usability, and interpretability are also paramount to the way in which the learner model may be appropriately opened to the learner, and also potentially limiting factors in terms of visualisation (Liang et al., 2012; Uther and Kay, 2003).

From this perspective, the level of detail available to the learner though the learner model is of key importance, together with the other types of information that may be represented concurrently. From the task driven perspective, this highlights a practical consideration of scalability, and informational sufficiency in different (potentially unanticipated) usage contexts. For this reason, the type of information and the approaches for structuring and facilitating its access are critical.

- Learner models can be opened in a rich variety of forms, revealing a great deal of information, not just about student knowledge. Sometimes more than just the core content of the model is externalised.
- Graphical and textual are the most common forms of visualisation of the learner model, although many novel forms of visualisation have been proposed. The method for externalising the content of the learner model is chosen to best suit the purpose and context of use.
- The visual form of the learner model should support the key aims for opening the model to the learner, which centre on promoting reflection and aspects of metacognition.
- Information structure is of key importance and the visualisation should include initiation points for metacognitive activity.

Chapter 2 highlighted that in the context of educational theory there are particular ways of supporting metacognition, including through formative assessment opportunities that present learners with access to information about themselves and their learning. This chapter has highlighted that OLMs are designed around these principles, that their externalised form is most commonly textual or graphical and that information structure is of key consideration. Visualisation methods support overviewing and access to more complex aspects of the underlying data, in addition to some making explicit aspects of the domain/underlying evidence in their representation. As such, this thesis next considers information that this technology makes available to the learner that is not directly part of their learner model but related to comprehension as an aspect important to the regulation of cognition (Section 3.6).

3.6 Open Domain Model and Learner Model Process Content

In addition to the direct opening of the learner model, there may be situations where it is educationally beneficial to open other related elements, such as the process by which the information is derived (e.g. Johnson et al., 2013a, Van Labeke et al., 2007) or corresponding information in the domain model (Johnson and Bull, 2009). Indeed, revealing domain content or the learner model process (see Section 3.2) may be key to interactive maintenance processes (Dimitrova and Brna, 2016; Ginon et al., 2016), or to describing or visualising misconceptions to the student (Johan and Bull, 2009). Arguably it is also key to acknowledge that some methods of opening the learner model also contain domain related information, even if this is just to show relationships between concepts (Zapata-Rivera and Greer, 2004; Johnson et al., 2013a; Mabbott and Bull, 2004). The domain information implicit in the OLM visualisation is likely to be of a coarser grain size than for specific articles of domain information, for example where structural information is the content shown. There may be situations where information from the domain is not fully known or is contributory to the modelling process e.g. where it is built into the design of multiple choice questions, which generate inferences for the learner model, but there is no domain model (Mabbott and Bull, 2004) or where inferences are generated by 3rd party educational systems (Kickmeier-Rust and Albert, 2016). In the case where learner models are built, for example, from a constraint-based approach (Mitrovic and Ohlsson, 2015), or the content of the models are specific articles of knowledge, this information is more readily available. Overall, an open representation of the domain model provides the context in which to interpret the information in the learner model in a directly comparable manner. Inspection of expert knowledge in this way might be considered as that of a glass box approach (Goldstein, 1982, in Paiva et al., 1995), or as an example of a successful learner against which the student may measure performance (Davis et al., 2016), or as an item considered as support to use in the SMILI open learner model framework (Bull and Kay, 2016). In practice this may only be done on finer levels of granularity, i.e. if the learner model is showing the level of knowledge (coarse grain), the equivalent domain model will always show 100% correct, and be of limited semantic value.

Returning to the educational benefits of a formative assessment approach outlined in Chapter 2, these can be considered from the perspective of facilitating reflection and metacognition. It is perhaps important to encourage learners to compare their actions to experts (Goodman 1998), to provide examples of strong and weak work (Chappuis, 2009; Shute, 2008) and to clarify what is good performance and why it is good, in order to provide opportunities to reduce the discrepancy between current and desired performance (Nichol and MacFarline-Dick, 2006).

- Part of opening the learner model might also show elements of the process by which information is arrived at, or elements of the domain content. This provides more contextual information which students can use to think about their learning (and their thinking).
- Information is on different levels of detail and contextual information may often only be done at finer levels of granularity.
- Additional contextual information is sometimes needed to make sense of the visualisation and may involve opening up or drilling down into the model for justification. This is highly relevant to supporting tasks relating to the regulation of cognition.

At this point, a recurring theme is that information exists at different levels of abstraction and different levels of granularity. Furthermore, the content of a learner model can be great and multifaceted. Techniques need to be employed to ensure that learner models can effectively externalise their content. This thesis therefore now considers these two issues in greater detail, looking first at information granularity (Section 3.7), and then at a drill down approach (Section 3.8), as a structural solution to support different grain sizes and issues of scalability. It then returns to consider these issues in the context of open learner models (Chapter 4).

3.7 Issues of Granularity

Information, digital or otherwise, manifests itself on a variety of levels of abstraction and often in large central stores of information (Busch, 2014). That is to say the grain size and level of aggregation of the information accessed may vary depending on a variety of factors, such as those related to the specific task to be performed, limitations of the data to be presented, the requirement for the information to be user-interpretable, issues of cognitive load, or limitations of the display paradigm or scalability (Lengler and Eppler, 2007; Bargiela and Pedrycz, 2002). It might also be further categorised in terms of whether it is semantic granularity (different levels of specification of an entity) or spatial granularity (resolution or representation at different scales) (Fonseca et al., 2002).

In defining what is meant by granularity, this metaphor considers grain sizes, whereby the smallest, most precise elements are fine grained and those which are at a higher level of abstraction (a lesser precision, or a lesser detail) are coarse grained (Khan et al., 2014). Thus, granularity is the extent to which an entity is broken down into its constituent parts, a level of precision, a level of indeterminability, or a level of aggregation (Bargiela and Pedrycz, 2002).

This also encompasses the size to which data can be sub-divided. Being able to access and locate key information quickly is an important aspect of this, and largely depends on the task in hand at a given moment in time (Khan et al., 2014).

Granularity also implies some hierarchical information in the level of abstraction of the information (Bargiela and Pedrycz, 2002), and can involve concepts such as subgrouping (for example in menu structures e.g. Wallace et al., 1987). In situations where datasets are large, it may not be possible to view all information at any one given time, and so some obscuration and aggregation may be required (Pedrycz, 2012). Where information may be of a finer granularity, there could be some effort required here to initially create, and subsequently maintain the information (Fonseca et al., 2002). Where user perception of information is required, information of different types may need some level of comprehension and visual categorisation during interpretation, and excessive amount of information can introduce a level of fatigue (Abrami et al., 2013), cause frustration (Pintrich, 2003), or add to cognitive load (Jacoby, 1977). Within technology solutions there are a variety of approaches to allowing information of different levels of abstraction to co-exist (see Lengler and Eppler (2007) for an overview).

Rationales for having information at different levels of granularity or abstraction range from supporting both browse and search and being able to access key information quickly (Khan et al., 2014), to being able to accurately ascertain a piece of information to a desired level of precision (Busch, 2014). These are task driven actions, in the context of self-directed learning that are relevant to the regulation of cognition, and also help support making information manageable, so that processes such as reflection may be balanced in terms of cognitive load.

Interactions may be driven by short and long term tasks or goals of students, and an informational need at the point of interaction, even if the student is not aware of what this is (see Chapter 2 for further commentary). Key issues for consideration here include:

- The task at the moment of interaction. Does this require getting key information quickly?
- Is the task at the point of interaction one of searching or browsing?
- Does the task consistently need detailed information, or consistently need high level information?
- Can all information be intelligibly displayed in the chosen visual method? How much information at the same level of abstraction can be presented before the visualisation becomes not fit for purpose, e.g. complexity, cognitive load, ability to perform task?
- As a dataset grows, is the way of viewing it scalable?
- Is there a natural structure to the data that should be presented?
- Is all available information to the same level of precision?

- Granularity may be considered as semantic (levels of specification) or spatial (part of a scale), and therefore can relate to levels of abstraction or resolution of the information.
- Granularity may vary depending on task, data limitations or issues of accessibility (such as cognitive load or interpretability). This is important in the regulation of cognition and supporting reflection at different levels of abstraction.

Coarser granularity information is most often presented first, and secondary to this is extra detail to support sub-tasks beyond a general overview. This supports the idea of information on demand and a level of accessibility, balancing issues such as cognitive load with the user's right to see any relevant information that is held. In a context such as self-directed learning it is important to consider that different informational types are required for supporting different processes, which in turn have potential to promote metacognitive activity. As such, approaches for structuring and navigating this are potentially highly relevant to the design of an open learner model that can support this (i.e. through student reflection and tasks relating to interpretation and actioning of information). Following on from this, this thesis next looks at the drill down approach in further detail, and then how this applies to the design of open learner models.

3.8 Using a Drill Down Approach

One approach taken, partially addressing the above issues of consideration, is that of being able to 'drill down' into a data set from a higher level. This is a trade-off between having full access to all detailed aspects of the information and benefiting from a reduction in cognitive load at the point of entry (Abrami et al. 2013), allows access to key information quickly (Khan et al., 2014), and follows the approach of overview first, followed by details on demand (Shneiderman, 1996). This avoids the ultimate loss of precision that otherwise exists from sole aggregation of fine grained information into a coarser grained representation, although some aspects of information being 'lost in translation' is inevitable (Busch, 2014). However, depending on short term goal, it is sometimes not desirable to view the information in its underlying form, where domains are large and complex, as it may limit usefulness (Bull and Kay, 2013). The use of the drill down technique is particularly driven by there being too much information to display at once in full detail, and cases where not all the detail is relevant all the time (Abrami et al., 2013). In almost all cases, a drill down approach is from information at a higher level of abstraction to more fine grained information (Buse and Zimmerman, 2012). Furthermore, the ability to drill into the data may be required from different perspectives, e.g. time, organisational structure, or architecture (Buse and Zimmermann, 2012).

The use of a drill down approach allows for an easy overview inspection of all aspects of the information to be visualised, with entry points to go deeper into this, permitting the user to get key information quickly: a particular challenge for larger data sets (Khan et al., 2014; Busch, 2014). The drill down approach is common in different technology types such as medical information systems (Aronson et al., 2016), computer systems (Norman, 1991), content management systems (Kaupp et al., 2013), search engines (Rosenfeld and Morville, 2002) and a variety of web-based services such as online banking, among many other examples. As might be expected, the consideration of supporting the use of a drill down approach may not be an issue until the scalability of the information set becomes an issue (Soo Yi et al., 2007).

Using this approach, most access to further detail occurs as and when an information need arises. Drilling down further into data may occur when tasks are expected to take longer to complete, or to get further information from a region of visual interest (Abrami et al., 2013) – which does also suggest that the presentation of the coarser grained information is highly influential, and potentially more so when browsing, as opposed to searching a dataset.

Furthermore, a common use of this approach is in fault finding, correcting problems, understanding problems, and identifying errors and deviations, or in gaining further insight and deeper understanding (Ikeda et al., 2012; Busch, 2014). These are arguably core elements of self-directed learning, of cognition and of metacognition, by supporting learners in reflective thinking through active navigational decisions (Abrami et al., 2013). Using a drill down approach for this purpose involves a process of independent working, and a series of cross-disciplinary skills of enquiry (21st Century skills, core to the concept of assessment for learning) and an element of reflection or consideration to identify discrepancies. If students employ this approach and the information being considered here is that of their own learner models then this arguably becomes an aspect of self-consideration, self-assessment, self-enquiry, or self-assessment. This also links to procedural and conditional aspects of interaction, and processes relating to when, why and how to engage with information for tasks relating to the regulation of cognition, or to inform reflective thought. These are key aspects of metacognition that self-directed learning and OLMs aim to support (see Section 3.3). This approach may also have further potential to mirror the level on which the learner is processing information, inferencing and sense making.

From this it can be considered that:

- the actions to drill down into data may be driven by task, by information representation, or by information volume.
- the approach is particularly situated to situations involving tasks of fault finding, correcting problems, comprehension and identification of errors and deviations.
- a drill down approach may support different behaviours, such as search and browse, and has potential to mirror the level on which the learner might process information.
- this aligns with approaches for supporting metacognitive activity.

- The use of a drill down approach has the potential to structure content in a form that is designed to support cognitive processing of the information. Information in focus is intended to better align with users' current tasks and processes.
- The use of a drill down approach is of particular use for tasks such as fault finding, problem solving, gaining insight and gaining understanding, which makes it of particular relevance to activities such learner reflection and self-assessment, and the regulation of cognition, which are core to metacognition. As such, drill down is an important aspect of an intervention designed to consider the impact of technology on metacognition.

This section has highlighted that a drill down approach (mantra: overview first, details on demand) is potentially a suitable structural solution for allowing learners access to information about their learning, regarding supporting interaction that may promote metacognitive activity. With particular reference to supporting declarative, procedural and conditional aspects of tasks such as comprehension, important in the regulation of cognition, it is relevant to the design of any technology that aims to facilitate this. This thesis next briefly explores the use of a drill down approach with other technology, and then in the specific use-case of examples of open learner models that implement this structural interface method.

3.9 Technology Supporting a Drill Down Approach

Some visualisation types support an element of drill down within them, such as fish-eye based approaches, which zoom in on a given part of the dataset for more detailed inspection (Tominski et al., 2006). Other technology solutions are also able to support this, such as tool-tips when hovering over items or inspection panes (Hearst, 2006), or a modal dialogue that populate with the content upon mouse interaction (Carey et al., 2000), and in many cases breadcrumbs are used to show place within the structure (Hearst, 2006). Menu structures, such as those popular in Windows-based systems in the 1990s and 2000s also reflect this

(Norman, 1991): categorisation of canonical components (e.g. applications) into a hierarchical structure of (sometimes) manageable component size, i.e. segmentation of the data (Wallace et al., 1993). In all examples, this is very much 'information on request'. In contrast to some educational technology systems such as intelligent tutors (see Section 3.1) that may provide more detailed information to the learner as and when it is considered appropriate (system initiated), the drill down approach is very much in the hands of the learner (user initiated), and is very much in alignment with open learner models, which aim to promote learner autonomy through this means (see Section 3.3).

Higher levels of granularity can also sometimes be used to provide an element of domain independence in terms of visual representation (as seen in some open learner models e.g. Mabbott and Bull, 2004). This can in some respects bring about a level of consistency where this is in part a requirement of the visualisation.

- Drill down approaches can be supported both through visualisation and interactive navigation. This provides information in a timely manner aligned with current workflow and information need, including some aspects of information granularity.

Chapter 2 highlighted that there are specific types of tasks important to the regulation of cognition, such as problem solving and debugging, and Sections 3.7 and 3.8 describe that information granularity and a drill down approach are highly relevant executing these, particularly so as the use of drill down technology is reliant on an active navigational approach by the user, which forces decisions to be made. This further indicates this is an important aspect of an open learner model both in terms of visualisation and interactive navigation.

3.10 Summary

This chapter has overviewed open learner models as a technology designed specifically to promote metacognitive activity though adaptively presenting the current state of student understanding back to the learner in a user interpretable form. Literature on common aspects of open learner modelling further highlights that information structure in terms of granularity and the ability to drill into it is highly relevant to aspects of metacognition (particularly in terms of performing key tasks surrounding data interpretation, where the data relates to identifying the student's current educational needs). These issues are prerequisite to making the formative assessment content accessible to the student. As highlighted in Sections 3.5 and 3.6 this implies that some level of abstraction or transformation may be required of the content (whether visual or informational), and may also require supporting content (e.g. from the domain) in order to make full sense of the formative assessment information that is afforded to the learner. It is therefore possible to consider that the use of a structural method (such as drill down) is itself a secondary aspect of the design of the technology in order to allow cognition to occur at the relevant level of abstraction and as to potentially better support some of the tasks Chapter 2 identifies as important in the regulation of cognition. To elaborate the implications of using these aspects together, Chapter 4 now considers existing ways that drill down methodologies are used in open learner models, and the issues and assumptions arising from this.

This chapter therefore contributes requirements towards the rest of the work in this thesis, with reference to the visualisation of the learner model and the implementation of a drill down approach.

Table 2: requirements.

ID	Description/Rationale	Supporting Sections
R4	Different visual/informational forms are needed, tailored to specific (relevant) cognitive tasks. Graphical and textual forms are most common, and this allows the OLM to be consistent with existing implementations.	OLM Visualisation Granularity Learner Models
R5	Use of a drill down approach allows for supporting tasks and capturing information relating to the regulation of cognition, and also a way for learners to access contextual information required to interpret visualisations. An evaluation of the consistency of use of drill down will be required to describe how it applies to OLMs.	OLM Visualisation Granularity Drill Down Open Domain Model

The thesis now moves on to consider how a drill down approach is currently facilitated in open learner models (Chapter 4).

Chapter 4

DRILL DOWN APPROACHES IN OPEN LEARNER MODELS

Chapter 3 highlighted ways in which OLMs show learners the current state of their understanding, provide information important in the regulation of cognition, facilitate self-assessment and help learners identify their current learning needs. It highlighted that information structure is an important aspect for promoting metacognitive activity, and that information granularity and drill down approaches are an important aspect of this. This chapter looks at specific examples of how information structure exists in open learner models, with reference to a drill down approach.

4.1 Open Learner Models, Information Access and Granularity

As stated in Section 3.5 , learner models may be opened in variety of different forms, with different visual properties. The SMILI framework (Bull and Kay, 2016) provides a common way to describe learner models. It postulates some variance in terms of granularity, stating that the learner model may be opened as knowledge level, knowledge, difficulties, misconceptions, and other elements including learning issues and preferences, depending on the reasons for opening the learner model to a particular stakeholder group.

To place these in some order of level of detail these might be considered as in Table 3, as a non-definitive suggestion. This ranges from a high level visualisation of knowledge level at the end of a coarser grain, moving through a decomposition of this into specific sub-categories of the knowledge level, and into a more detailed description of what the specific articles of knowledge might be. An even finer granularity of information may consider the origins of specific beliefs and aspects of epistemology (see Johnson et al., 2011; Johnson and Bull, 2011 for more commentary). This is presented as a means to suggest that the granularity of information that is available to be presented in a learner model has great variance and is potentially a very rich tapestry of information.

Table 3: granularity of learner model externalisation.

Coarse /\\ | | | | | | | | | | \/ Fine	• General information, general goals, aggregated information. • Level related information; i.e. ability to correctly apply a specific rule; whether something is known; aggregation of other framework items, or partial model. At this granularity, visual methods that quantise the OLM are coarser than those that use continuous scales. • Level related information broken down into sub-nodes (e.g. as per a domain structure) or other information, such as structural relationships. • Level related information, justified in terms of other model information such as specific model weightings, evidence volumes, and/or quantification of the modelling process. • Examples of applied understanding, specific learner beliefs and inferences. E.g. specific rule applied. • Examples of applied understanding broken into sub-components. • Justified applied learner model beliefs in terms of other model information, epistemic backing, and/or modelling process description showing explicit link between evidence and modelling outcome.

Building on the idea of information on demand, identified as important to a drill down approach in Section 3.8 and to establish the extent to which different information types are concurrently presented in existing OLM systems, this thesis next reviews a range of OLM systems from the last two decades (Table 4).

Table 4: open learner models opened for students.

System	Use of a drill down approach	OLM on different levels of granularity	Multiple visualisation methods	Access of one visualisation or information type through another	Expansion of contraction of domain structure to amend the level of detail shown	Inspection of the modelling process	Inspection of domain content (not including domain information to label or structure information)
AniMis (Johan and Bull, 2009)	Yes	Yes – misconceptions only	Yes	Yes	No	No	Yes – misconceptions only
CALMsystem (Kerly et al., 2008)	No	No	Yes	No	No	Chatbot can provide justifications when challenged	No
CosyQTI (Lazarinis and Retalis, 2007)	No	No	Yes	No	No	No	No
COMOV (Perez-Marin, 2007)	No	No	Yes	No	No	No	No
C-POLMILE (Bull et al., 2003)	No	No	Yes	No	No	No	No
EER-Tutor (Mathews et al. 2012)	No	No – although some visualisations show domain structure	Yes	No	No	No	No
EI-OSM (Zapata-Rivera et al., 2007)	Yes	No	No	Yes – to get to evidence layer	No	Partial – through an evidence based approach	Partial – links to associated applied domain content
E-KERMIT (Hartley and Mitrovic, 2001)	Yes	Yes	No	Yes	Yes – tell me more	No	No
ELM-ART (Weber and Brusilovsky, 2001)	Yes	No	Yes	No	Yes	No	No
Flexi-OLM (Mabbott, 2009)	No	Yes	Yes	No	No	No	No
Haptic Learner Model (Lloyd and Bull, 2006)	No	No	Yes	No	No	No	No

System	Use of a drill down approach	OLM on different levels of granularity	Multiple visualisation methods	Access of one visualisation or information type through another	Expansion of contraction of domain structure to amend the level of detail shown	Inspection of the modelling process	Inspection of domain content (not including domain information to label or structure information)
INSPIRE (Papanikolaou et al., 2003)	Yes	Yes	No	No	No	Yes – information button explains how knowledge level is evaluated	No
LEA's BOX OLM (Ginon et al., 2016)	Yes	No	Yes	No	Yes – retract - use of filters	Yes – as part of negotiation dialogue	No
MusicaLM (Johnson and Bull, 2009)	Yes	Yes	Yes	Yes	Yes – expand - select from list	No	Yes
Mr Collins (Bull and Pain, 1995)	No	Yes – dialogue statements of different types	No	No	No	No – but evidence can be presented	No
MyExperiences (Kump et al, 2012)	Yes	No	No	No	Yes – retract - use of search – expand – click of treemap item	No	No
NEXT-TELL OLM (Johnson et al., 2013b)	Yes	No	Yes	Yes – for modelling process	Yes – retract - use of filters	Yes	No
OLMlets (Mabbott and Bull, 2004)	Yes	Yes – misconceptions only	Yes	No	No	No	No
QuizMap (Brusilovsky et al., 2011)	Yes	Yes	No	Yes – mouse interaction	Yes – retract – mouse interaction	No	No
STyLE-OLM (Dimitrova, 2003)	Yes	Yes	Yes	Yes	No	No	No

System	Use of a drill down approach	OLM on different levels of granularity	Multiple visualisation methods	Access of one visualisation or information type through another	Expansion of contraction of domain structure to amend the level of detail shown	Inspection of the modelling process	Inspection of domain content (not including domain information to label or structure information)
Subtraction Master (Bull and McKay, 2004)	No	No	No	No	No	No	No
SQL-Tutor (Mitrovic and Martin, 2007)	Yes	No	Yes	Yes – tell me more	Yes – expand – use of tree nodes	Yes – collapsible tree structure	
SIV (Kay and Lum, 2005)	Yes	Yes	Yes	Yes – infer, show more	Yes – expand, retract, search	No, but information about it is available	No
t-OLM (Ahmad and Bull, 2008)	No	No	Yes	No	No	No	No
TAGUS (Paiva et al., 1995)	No	Yes	No	No	No	No	No
UM toolkit (Kay, 1994)	Yes	Yes	No	Yes – interaction with nodes	Yes – expand - interaction with nodes	No	No
UMPTEEN (Bull et al., 2007)	No	Yes – misconceptions only	Yes	No	No	No	No
VisMod (Zapata-Rivera and Greer, 2004)	Yes	Yes	Yes, customisable	Yes	Yes	No	No
xOLM (Van Labeke et al., 2007)	Yes	Yes	Yes	Yes – tell me more	Yes – expand - select from list	Yes – Toulmin based justification	No

Table 4 emphasises the range of approaches that can be taken when externalising the underlying model. Some employ interface techniques such as to have a main summary as an entry point in the model with modifiers to amend the scope of this, e.g.: allowing searching or filtering (e.g. Kump et al, 2012); expanding elements of information focus to reveal sub-components (e.g. Hartley and Mitrovic, 2001); or using zoom-like functionalities to increase or decrease the volume of content shown (e.g. Kay and Lum, 2005). This is in line with Shneiderman's (1996) guidelines on information presentation (see Section 3.8). Examples also exist whereby the selection of an informational focus presents a greater detail about that particular aspect only (e.g. Mitrovic and Martin, 2007) i.e. inspection of sub-components, a finer granularity, but in the same visualised form as for the coarser content. Further examples exist where this results in finer grained information but of a different form (e.g. Van Labeke et al., 2007). These approaches overall might be considered as a 'drill down' process. Of the examples that visualise the entirety of the learner model at once (at the level of granularity upon which they are opened, even if the underlying model may be more complex), these tend to be with smaller volumes of information and higher levels of descriptions (e.g. Brusilovsky et al., 2011).

There are examples along the lines of Mabbott and Bull's (2004) work advocating multiple representations of the learner model (see Bull and Kay, 2013 for review). In cases such as these, the different visual forms mandate slightly different aspects of the information (or structural information) around it to be included to allow it to be interpreted – for example a concept map may be considered as a slightly finer grained representation than a skill meter, as it contains additional information about relationships that exist between nodes, in the context of the domain. Switching between these visualisations may thus also be considered a change in granularity, even if the main content is comparable (e.g. knowledge level only).

Of those learner models where, upon the selection of an informational focus, further information is presented, this may be of a more intricate level of granular detail, beyond breaking down into sub-components. It may give applied examples (Kay, 1994), specific misconceptions (Mabbott and Bull, 2004), or epistemically-based descriptions justifying the learner model (Zapata-Rivera and Greer, 2004).

Those learner models that go into further detail using this drill down approach often do this for the purpose of opening up aspects of the learner model interpretation process (Van Labeke et al., 2007), and in some instances make links to the underlying evidence layer (Johnson et al., 2013b) where this is present. Several examples also forge links between the OLM and corresponding content from the domain model (Johnson et al., 2009), related domain content, suggested resources or other formative assessment opportunities (Zapata-Rivera and Greer, 2004).

With those systems that employ an element of interactive maintenance or are reliant on dialogue-based interaction (e.g. Kerly et al., 2008; Dimitrova, 2003; Paiva et al., 1995) there is much that is initially hidden from the student, in terms of the precise form in which the model is opened to the learner, be it of a textual (e.g. Bull and Pain, 1995) or graphical (e.g. Dimitrova and Brna, 2016) form. The argumentation constructed may be to elaborate, clarify or justify student understanding, and is very much taken from the approach of 'information on demand'. Again potentially because of the complexity, volume, or finer granularity of the information, this paradigm and level of interactivity is selected as it best suits the aims of the system and/or educational approach. Its dynamic and reactive nature also implies that it cannot be constructed or anticipated in a finite or fixed structure but is instead requires a workflow. Arguably, in many cases this is a drill down into the system or student understanding, with a view to supporting student reflection and improving the quality of the model (Zapata-Rivera et al., 2007).

- Learner models are opened to the learner with a view to promoting metacognitive practice, and drill down approaches exist in some OLM systems to support access to the learner model (and domain or evidence) information, whether this is a formal access of a more detailed component through a more coarse grained one, a change of visualisation type, an interactive dialogue, or a level of interactivity to expand/contract the amount of the domain shown.
- The specific approach taken with drill down in OLMs is consistent with its use in other technologies (See Section 3.9). It is suggested that drill down is used to access information that is generally more specific, of finer granularity, or of a different information type. Drill down is particularly used where models are larger.

OLMs are intended to be designed to support metacognition and self-assessment. The way in which they present information is important when balancing a cognitive load overhead such that metacognitive practice may take place and drill down is one method by which this may be achieved. A review of OLM systems shows that aspects of a drill down approach (information on demand, increasing granularity, information structuring) are present in the design of some OLMs. This thesis next considers issues around the use of this specific approach to OLM design with several case studies.

4.2 Applied Examples of Information Access in OLMs

The OLM system case studies reviewed in this section give specific examples of how aspects of granularity and drill down approaches are implemented in open learner models, and the rationale behind the related design choices. The first study considered gives a good summary of the information presentation challenge that has existed from the outset of open learner modelling research, and it proposes using tree structures as a solution to manage information volumes. More specifically this section looks at the different types of information available in the OLM visualisation component, how it is structured/accessed and how use of drill down has been implemented as a practical solution to facilitate the aims of the technology. The 7 case studies are presented in their order of implementation, giving commentary on how this issue has been addressed across the lifetime of the field.

4.2.1 UM Toolkit (Kay, 1994)

Kay (1994) makes use of tree structures to represent and visually organise information in the learner model. The initial overview representation is achieved by changing the presentation of the node label (e.g. shape, colour, size). Leftmost items are of the coarsest granularity ('partial models' where they are aggregates of the models of sub-components). Interaction with nodes shows finer grained aspects of the model, not limited to the models of subcomponents, but showing articles of knowledge or applied beliefs. Interaction through this approach 'drills' further into the model to present information of a finer granularity. This approach allows visualisation of the model as a whole, a clarification of structure and support for navigation.

Kay (1994) argues that the model should be effectively accessible for all parts of interaction, and to do this it should strive for simplicity at all levels of granularity, in the size of the model and in the depth of the reasoning process. It is underlined that learners must be able to cope with both aspects of this. This approach favours the separation of elements within the reasoning process, so that students may limit the focus of their attention to the parts of the model that are relevant to a given situation or task.

Aspects of accessibility and understandability are both emphasised, in terms of the persistence of data, the defined relationship between the learner and the system, and the ability for the model to be useful and defensible in terms of the visualisation of the underlying model. It is suggested that a coarse grained model might be best used for more sophisticated or complex aspects, and a fine grained model for simpler ones. However, it is also fully acknowledged that while a summary might make a model more accessible, the learner may find it find it difficult to appreciate the full meaning using this approach. Indeed, as part of a process of enquiry, as might be best suited to a self-directed learning/reflection approach, occasions may exist where access to a partial model is all that is needed to answer a query or for students to establish whether they have an interest in its content.

The role of time is also important here. Kay describes the process of 'accretion', where the model has a steady flow of new evidence that is gathered about the student. It shows that the collection, interpretation and management of evidence about the learner is a central task for the modelling system, including compacting and disregarding non-relevant items, or acknowledging insufficient information earlier on during interaction. In addition to this, there are changes in the learner that will be reflected in the model, not limited to learning, forgetting, and changing preferences.

Each of these aspects are important in terms of the design of an open learner model. First, different ability/skill levels of students may have an effect on which parts of the model are relevant, accessible or interpretable; secondly there is no assumption that all parts of the model will be accessed, or interaction tasks completed consistently across interaction, based on students' individual differences and own self-directed learning strategies; and thirdly, where information is presented using a drill down approach to make it usable, it is not clearly anticipatable which aspects of the model are likely to consistently bring about the greatest educational effects.

4.2.2 STyLE-OLM (Dimitrova, 2003)

STyLE-OLM is designed around a level of interactivity with an element of interactive learner model maintenance, which aims to foster reflective thinking and increase the accuracy of information within the model. In contrast to the more static approach taken by Kay (1994), students can either browse through the content of their learner model or engage in interactive dialogue, for such purposes that might be categorised as exploratory (collecting more information), explanatory (discovering why an error exists) or negotiatitive (clarifying viewpoints). Interactive dialogue can be textual or graphical (conceptual graphs) and through this interactive process more detailed and fined grained information held in the learner model can be externalised, such as in the system justifying its position, thus controlling the granularity of information presented. Information in the learner model may be considered as correct, erroneous or incomplete.

This level of interactivity brings a further sense of openness, and learners' active engagement in navigational decisions has potential to allow learners to engage in reflective activity, such as where the interaction prompts learners to recall or reconsider ideas, validate their beliefs, revisit arguments or investigate evidence. This may be particularly the case where the learner's beliefs differ from those of an expert and where quick (overview) responses may be inappropriate or incomprehensive.

4.2.3 SIV (Kay and Lum, 2005)

In a similar underlying approach to Kay (1994), SIV also makes use of a tree-like structure as an entry point into information. With concern for the interrelatedness and size of the ontology that structures the information, fonts are used to show related concepts, horizontal positions for evidence volumes, and colour to show aspects of the learner model. This in part makes information accessible through elements of visual encoding that could otherwise be revealed as part of drilling into the information (details on demand) and this is more advanced than a basic tree structure. This is a simple overview from where, upon mouse interaction, further information about a specific aspect of the model is shown, as part of a drill down process. Interface options also exist for showing more or less detail in the OLM, and for it to be searched to narrow the scope of information. The drill down approach is used in the instance where the limitations of the semantic visual encoding are reached.

Kay and Lum argue that evidence from more fine grained inferences may be needed for more fine grained learning goals and endorse the argument that *"to make models useful for reflection, they must model the learners at varying levels of granularity: coarse learners see how well they're doing on overall goals, and fine grained can determine which elements of work contribute to high level goals."* Learners should be able to feel in control and able to delve into details and processes underlying the model. Also highlighted is that open learner models "*can easily provide very large amounts of data about a learner, although this can be of varying and poor quality*". Indeed, part of some coarser grained representations of the learner model may have no direct evidence, however, students can appreciate the level of reasoning required for this. Of the finer grained elements, imports to the system, such as those from quiz results, class exercises or multiple choice questions, exams, tutorials, assignments or projects may be able to give more fine grained information as these refer to a small aspect of domain content. The role of the system is to give control to the student for metacognitive purposes and even simple explanations of aspects of modelling and the underlying model might be effective.

4.2.4 CosyQTI (Lazarinis and Retalis, 2007)

As in the above case studies CosyQTI uses different visual methods in conjunction to support information inspection. There is a simple bar graph and a slightly finer grained textual data representation, both of which are combined with activity-based learning analytic information. Learners can see their progress both per assessment and per section, depending on what is needed to be known to identify further areas of study. Highlighted are occasions when more detailed information is needed, including where students are overconfident, where misconceptions exist and where there are discrepancies between different classifications applied – e.g. the understanding of what is classified as “good” may vary between students (and educators) and from context to context. Learners also asked for more precise reports when problems were apparent, particularly with reference to evidence and inferences – what is wrong or unanswered, and with the right answer shown for comparison (application of domain content). This is important in tasks relating to the regulation of cognition, such as debugging and comprehension.

Also argued is the need for frequent access in terms of feedback in order to understand progress and issues, which is a common approach and is support that open learner models can provide in this formative assessment and self-directed learning context. Indeed, Lazarinis and Retalis state that undertaking short amounts of testing before model inspection is important with particular reference to understanding the adaptive nature of the open learner model, and that it is important for students to find a link between the information they view and their learning processes, such as would be a key aspect of metacognition.

4.2.5 EI-OSM (Zapata-Rivera et al., 2007)

A further example from the same time period is that of EI-OSM, which has a more graphically enriched interface that requires students to take much more of an active navigational approach. The OLM is a proficiency map, where colour is used to indicate the knowledge level of the student. The links that exist between nodes reflect the structure of the domain (i.e. sub skills), thus the main presentation of the learner model shows different levels of granularity concurrently. Through mouse interaction the learner is able to go one level further and inspect the evidence layer, which shows supporting content that includes details of students' actual answers. Challenges to the model and showing evidence are each considered from the perspectives of credibility, relevance and quality. A Toulmin-type argumentation is used to substantiate the validity of information that is shown, with a view to building critical thinking and reflection skills, such as are highly relevant to metacognition.

The modularity of information presented to the student is intended to support a self-paced approach, and providing links to materials supports the student in their planning with regard to what needs to be done to correct an identified problem. The system uses an 'active reports' approach to show relevant evidence to the learner when the model is queried, and more information is requested on demand. This can show different aspects of the learner model, information on different levels of granularity, different external representations and changes in light of new evidence. This is in part to address the viewpoint that students often see a learner model as something "complex, obscure" and an external entity that is not always trusted.

An important point is made with regard to the sources of information for the model that underlines that not all evidence is to the same validity, reliability, accessibility or level of granularity, thus allowing perception of the limitations of the information as part of a conscious cognitive processes. This is a key challenge in a multi-source environment. Often the level to which the information can be modelled with any certainty is the coarsest common denominator.

4.2.6 xOLM (Van Labeke et al., 2007)

In a similar way to which EI-OSM implemented a Toulmin-type argumentation approach, xOLM also aims to go beyond the black box approach, giving additional learner model information about how the modelling has taken place and is justified, also using Toulmin-like argumentation. It considers ways that the learner can benefit from engaging with this level of information in terms of how learners think about their learning. It is argued that the important aspects of this are the *summary* (aggregated model of the information) and the *facts* (such as those identified, inferred or diagnosed during interaction with the system). However, there can be a wealth of intermediate information to connect the two, which can be complex and great in size, and which might need a level of reorganisation in order for the model to clarify any justifications to the learner. Indeed, one significant limitation to a learner model of this type is the sheer volume of evidence that could potentially be gathered. Learners will not be able to assimilate, access and interpret everything at once, and therefore an interaction method is required in the model to make information available on demand, or upon query. From a high level list of domain content and representation of the learner model the learner can drill down further to obtain more information, justifying the representation in the learner model and making explicit the link between this and evidence that underpins the representation. An emphasis is placed on the need for different aspects of the learner model to be opened in different ways, particularly where there are questions about making transparent the accuracy of the underlying model.

Depending upon the nature of the informational query or demand of the learner (as might originate in a self-directed learning or learner reflection scenario) there will be different informational needs. Some of these may be intensive and some more fine grained than others. It is therefore only at certain points during interaction that more detailed justification of the learner model is needed, that nodes are required to be expanded and that claims should be justified with sub-claims. Challenges or inspections of the model may, for example, indicate disagreement with the information, a desire to validate understanding, or be sense-making to further comprehend what is understood and why.

The underlying assumption of this approach is that learners will be directed towards the actual state of their beliefs, rather than toward the trajectory of their abilities, as is aggregated at higher levels of granularity. Finding a suitable interface approach to support a drill down or justification-based paradigm is an upcoming challenge, as it needs to support both finding appropriate information in the learner model relevant to current goals/tasks/information needs, and at the same time should clearly relate this to the overall picture of the learner given by the system; i.e. the link between coarse and fine grained information must not be lost. This might be achieved by having both elements together in juxtaposition. A less complex representation of the learner model is more accessible; however, this is a loss of precision as compared to what could potentially be presented.

Information with which the learner model is built may likewise fall into two categories: behavioural (such as performance in an exercise) and diagnostic (such as self-reporting of affect, such as confidence). With reference to motivational factors, such as learners' confidence in their ability to correctly answer questions on an element of the domain, diagnostic information is obtained via self-report and self-assessment, whereas behavioural evidence comes from the student's interaction with educational material within the learning environment. The alignment of the two would indicate that the student's perception of their ability matches their ability to apply understanding, and so is a measure of self-assessment accuracy.

4.2.7 MyExperiences (Kump et al., 2012)

The final case study looks at a further example of graphical hierarchical structuring of data using more state-of-the art visualisation methods, with the same underlying principles as the tree structure approach of Kay (1994). In MyExperiences learners inspected their model using treemap structures and this was an implementation of some of Shneiderman's (1996) guidelines including initially presenting the user with an overview, after which details can be accessed on demand. In the OLM, the overview could be reduced in size (searched) or particular areas of interest expanded, taking more of an interactive approach. All information related largely to knowledge level and did not go into the level of detail of applied understanding (spatial granularity, rather than semantic granularity). Highlighted in this system is the importance of students being presented with an overview, and also problems of scalability, particularly with large datasets, which is a pertinent problem for newer OLMs to address as electronic data becomes more pervasive.

Furthermore, with MyExperiences, Kump et al. also highlight that complex inferences are not necessarily understandable by learners, nor can some visual methods easily support the learner to the same extent when going deeper into the OLM data if additional elements of complexity such as model weightings are required to be visualised. Similarly, where additional detail is accessed this might indicate looking for hints for a task or an attempt to acquire knowledge relating to a concept. Overall, being able to drill into the OLM in more detail gives opportunity to enhance transparency and understandability, and hint towards aspects of uncertainty in the data, but additional information from the domain may need to be presented to help interpret finer grained elements of the model. From the perspective of supporting metacognitive tasks, the interface approach supports students in terms of self-assessment and reflection by making the information accessible/interpretable and presents decision points in terms of interactivity or an enquiry process, relevant to tasks that relate to the regulation of cognition.

4.3 Summary

Open learner models are designed to promote learner reflection, self-assessment and metacognitive practice. At the design stage many implement a drill down approach to structure information (either replacing content (e.g. Van Labeke et al., 2007), or supplementing content, e.g. (Kay, 1994, Johnson et al., 2013b)) to aid aspects of metacognitive practice. This is often done where natural limits are reached in terms of semantic visual encoding of information, both in terms of scale and of information type. *It is common to use drill down as a structural solution to interface design in OLMs and educational arguments are made for informally ordering information in this way, although no studies have looked directly at the use of drill down in OLMs in detail.*

In many of the examples surveyed information is split so that it is accessible for relevant cognitive tasks, while leaving the choice of task and reason for interaction up to the learner. This leaves learners to foster their own strategy for engagement as best suits their self-directed learning approach. Information is separated in a way that mirrors different parts of the reasoning process. This builds on the idea that learners need to have a level of information that allows them to feel in control to further effective engagement with the OLM, for reflective purposes or for the completion of tasks relevant to the regulation of cognition. Additionally, arguments are made that students should be provided with high level overviews of their model (Bull and Kay, 2007, Uther and Kay, 2003). This is to provide support for interpreting the information, such that students can connect summaries to more detailed facts about their understanding and validate what is understood and why. As stated in Chapter 3 and in several of the case studies, this approach is important to support problem solving and learner reflection as mirroring cognitive structure provides appropriate opportunities (formative assessment interventions) for students to think about their learning and how they are learning. It is well acknowledged that the precise nature of the tasks the student is attempting to undertake are not always well known at the time, and sometimes not fully known by the learner. *From the OLM examples, and the literature, this thesis considers that splitting information on the interface according to relevant cognitive task is an important aspect of design for learner model tools designed to promote metacognition and self-assessment. This fits well with a drill down approach.*

Aligning with the idea of self-directed learning, and the custom nature in which students may access extra information for the purpose of thinking about their learning (as established in Chapter 2), there is no assumption that the learner should/would make use of drill down as a matter of course, and likewise drilling deeper into the information does not necessarily mean that they have little knowledge of the concept (Kump et al., 2012). Literature has also shown that ineffective learners may struggle with more detailed aspects, and furthermore even simple model representations can have positive learning benefits for students. *It will therefore be important to first understand student acceptance of the drill down approach and open learner model.*

From the survey of systems (Chapter 4), OLM literature (Chapter 3), and that of self-directed learning in Chapter 2, the system designs detailed in this chapter make assumptions about the tasks for which students will use the software. This is something that cannot be fully regulated in systems where the focus is to change the internal representation of student understanding, and potentially the methods for interaction (as such might arise from metacognitive activity). It is also acknowledged that multiple strategies can lead to equally successful outcomes (Gobert et al., 2013) and even coarse representations can have positive effects, especially with weaker students (Mitrovic and Martin, 2007). *This thesis therefore considers that a behavioural description of how the system is being used is highly relevant to understanding its impact in areas such as self-assessment and the potential completion of tasks that are regulatory to cognition.*

Information to build the learner model may be behavioural (e.g. performance data) and diagnostic (e.g. self-reporting of affect e.g. confidence). As a measure of determining changes in self-assessment accuracy, alignment of the two indicates that the students' perception of their abilities may match their ability to apply understanding. *A change in the alignment may indicate a change in ability to accurately self-assessment,* which is important in contributing to metacognitive development.

From the systems surveyed, the provision of information at finer levels of granularity is also built on the assumption that this will direct learners towards reflecting on the actual state of their beliefs. This is in contrast to viewing information that relates only to the trajectory of their ability, as is shown by coarser grained or summary representations. This is underpinned by the definition of reflection as conscious exploration of experience and a method for making thinking processes explicit (see Chapter 2). The link between coarse grained and fine grained information must not be lost for it to remain genuinely useful to the learner for metacognitive purposes, and it may be used to allow access to reasoning processes, underlying data, origins of the data, or be part of interactive maintenance of the model. In terms of design this sometimes translates as producing more detailed reports and access to additional information when problems are apparent at surface level, when justification is required (e.g. using methods such as Toulmin argumentation), or when limitations of the information need to become apparent. Likewise, regular use of the technology, in between small updates, is important for forging the relationship between the content and their learning process. Overall these methods encourage learners to *engage with aspects of information that are required to promote reflective and metacognitive activity, through making active navigational decisions, under the learner's control. Patterns of interaction with these at different levels of abstraction may be used to indicate participation in tasks regulatory to cognition.*

- It is common to use drill down as a structural solution to interface design in OLMs for educational rationale, but no studies have looked directly at the use of drill down in detail.
- From the OLM examples and the wider literature, this thesis considers that splitting information on the interface according to relevant cognitive task is an important aspect of design for learner model tools that intent to promote metacognition. This fits well with a drill down approach.
- It will be important to first understand student acceptance of the drill down approach and open learner model.
- A behavioural description of how the system is being used is highly relevant to understanding its impact in areas such as self-assessment and the potential completion of tasks that are regulatory to cognition.

- A change in alignment of student perception of their abilities and their ability to directly apply understanding may be used as a measure of self-assessment accuracy.
- Overall OLMs encourage learners to engage in aspects of information that are required to promote reflective or metacognitive activity, through active navigational decisions that learners take under their own control. Patterns of interaction with these at different levels of abstraction may be used to indicate participation in tasks regulatory to cognition.

A consideration of existing applications of a drill down approach in OLMs contributes the following requirements to the evaluation of the research question:

Table 5: requirements.

ID	Description/Rationale	Supporting Sections
R6	Information should be split broadly according to cognitive task/granularity and transition between these should be supported by an active navigational approach. This will describe how drill down is used.	Information Access/Granularity Applied Examples in OLMs
R7	The system must maintain and record comparable models of behavioural (e.g. performance) and diagnostic (e.g. self-reporting) data. This will allow evaluation of changes in self-assessment accuracy and a temporal description of how learners are approaching the domain content.	Applied Examples in OLMs Discussion of Applied Examples
R8	User acceptance of the open learner model and drill down is required to contextualise any usage behaviour, it cannot be assumed.	Applied Examples in OLMs Discussion of Applied Examples

This thesis now moves to revisit the assumptions raised in Chapter 1 and summarises salient points in the literature to present sub-research questions.

4.4 Questions Arising from the Literature

Chapter 1 introduced several assumptions that the literature review aimed to address:

- **Aspects of metacognition is measurable:** Metacognition can be defined (Chapter 2) and comprises knowledge of cognition and regulation of cognition. Metacognition is very much an internalised process, and whilst changes in behaviour can indicate metacognitive activity, investigation of these aspects will need to be in tightly controlled experimental settings in the case of knowledge of cognition, and so in a self-directed learning setting this can be considered fairly intractable. The aspect of metacognition that relates to the regulation of cognition manifests in tasks such as problem solving, and so evidence of the completion of actions consistent with these tasks may more readily be used to indicate metacognitive activity of this type. Records of student interactions with OLMs may be used to capture suitable information to describe behavioural elements required to indicate this. (See Chapter 2 to Chapter 4.)

- **Mechanisms exist for measuring impact on self-assessment:** Self-assessment is an important regulatory process that contributes to metacognitive development. Aspects of this are measurable such as student ability to accurately self-assess. A suitable way to do this could be to keep to the same precision a model behavioural model (what the student did) and a diagnostic model (what the student thought that they did) and compare changes in the two across time. (See Chapter 2, Chapter 4.)
- **A drill down approach is suitable for an open learner model:** A review of examples of OLMs from the last two decades indicates that a drill down approach is often implemented as a structural solution to make information available to the student in manageable chunks, or on an appropriate level of abstraction as to best support a particular cognitive task. It is appropriate to apply this technique to an OLM and appropriate to investigate the impact of doing so. (Chapter 4)
- **Self-assessment and metacognition are areas where an open learner model should have impact:** A consistent narrative is presented in OLM research that the learner model is opened to the learner for metacognitive benefit and to support self-assessment practice (Chapter 3). The literature on education (Chapter 2) highlights that these are important and beneficial aspects of student development, and that technology, formative assessment and self-directed learning contribute to this. It would therefore be appropriate to expect that the use of an OLM in a self-directed learning setting should impact on metacognition and self-assessment.

With the initial assumptions clarified, this thesis now considers the research question in further detail:

The research question posed by the thesis is "**how is an open learner model with a drill down approach used and what is its impact on self-assessment and metacognition**?" As has been covered in the literature review of Chapter 2 to Chapter 4, it is appropriate to ask this question because:

- **A drill down approach** is suggested by the literature as appropriate to use in an open learner model for metacognitive and cognitive benefit. It is often implemented in the design of open learner models, but its use is seldom evaluated in detail. ***This is a gap in the literature.*** This leads to the sub-question ***"how is an open learner model with a drill down approach used?"*** which should be answered with an emphasis on user acceptance and consistency of use, as has been evidenced as a standard approach by other studies that have taken place in the field of open learner modelling (see Chapter 3, Chapter 4).
- **Self-assessment** practice contributes to metacognitive development and is a rationale for opening the learner model to the learner. Changes in self-assessment accuracy are observable through comparison of diagnostic and behavioural learner models, and it is valid to seek new settings in which student self-assessment may be improved. ***This has not been investigated with specific reference to a drill down approach.*** This leads to the sub-question ***"what is the impact of an open learner model with a drill down approach on self-assessment accuracy?"*** which should be answered with an emphasis on changes in accuracy across time, as has been the approach of other studies of student self-assessment practice with OLMs (e.g. Mitrovic and Martin, 2007; Kerly et al., 2008), and an additional emphasis of how this varies with drill down use, building on the gap in the literature.

- **Metacognitive benefit** is a reason for opening the learner model to the learner, and again it is valid to seek new settings in which metacognitive benefit may be achieved. Metacognition includes knowledge about cognition (which is very much an internalised process and difficult to measure) and also the regulation of cognition, which translates into tasks such as problem solving and comprehension, which have measurable components and strong educational benefits. As per the literature of Chapter 2, there are behavioural descriptions of actions such as problem solving processes (i.e. problem identification, careful inspection, execution of interaction strategies such as trial and error, and monitoring feedback on the success of actions). The presence of a drill down structure is designed to support this (Chapter 3) and makes elements of the behaviour explicit in student interaction. ***As evidence of OLMs being used for metacognitive benefit is an ongoing line of enquiry and drill down is designed to support aspects of tasks important in the regulation of cognition,*** this leads to the sub-question ***"what is the impact of an open learner model with a drill down approach on the support for the regulation of cognition?".*** The behavioural descriptions in Chapter 2 indicate that this should be answered with an emphasis of student behaviour consistent with different phases of problem solving, including ability to identify problems, careful inspection of problems as distinct from areas of student uncertainty, and evidence of strategies such as trial an error/hypothesis testing with regular inspection of the outcome of actions.

Building on this, and having considered the nature of metacognition and self-assessment (Chapter 2), and of open learner models and a drill down approach (Chapter 3 and Chapter 4), this thesis next presents a plan for an evaluation of the research question (Chapter 5). The implementation of technology suitable for investigating issues relating to drill down in an open learner model is presented in Chapter 6, and the evaluation of the research question with two studies is reported in Chapter 7 and Chapter 8.

Chapter 5

RESEARCH METHOD AND ETHICS

This thesis posed the question "*how is an open learner model with a drill down approach used and what is its impact on self-assessment and metacognition?*" The literature review has clarified initial assumptions about why it is appropriate to ask this and how it may be answered. This chapter presents the research method by first revisiting question and then details the participants, materials and methods required. The requirements identified in the literature review are revisited in Section 5.7 and ethical considerations are reported in Section 5.8 .

5.1 Summary of Sub-Research Questions

As summarised in Section 4.4 the research question breaks down into three smaller sub-questions, each with a proposed series of emphases on which the analysis should focus, as informed by the literature review. The evaluation will comprise the following analysis:

1) **How is an open learner model with a drill down approach used?**

   The focus is placed on user acceptance and consistency of use. The analysis will consider:

   a. Is an open learner model with a drill down approach accepted by its users?

   b. Is use of the drill down approach in the open learner model consistent across time?

   c. Is drill down always used when inspecting the open learner model?

2) **What is the impact of an open learner model with a drill down approach on self-assessment accuracy?**

   The focus is placed on changes in accuracy across time with drill down use. The analysis asks:

a. Will student self-assessment accuracy increase over time when using an open learner model with a drill down approach?
b. Will student self-assessment accuracy increase to a greater extent when greater use is made of drill down?

3) **What is the impact of an open learner model with a drill down approach on the support for the regulation of cognition?**

   The focus is on student ability to identify problems, careful inspection of problems as distinct from areas of student uncertainty (evidenced through the use of drill down), and evidence of strategies such as trial an error/hypothesis testing with regular inspection of the outcome of actions (regular inspection of domain content in one area at a time). The analysis asks:

   a. Will students use drill down in the OLM to inspect information about problems?
   b. Will students use drill down in the OLM to inspect information about areas of uncertainty?
   c. Will students use an OLM with a drill down approach to focus on one domain area at a time?

5.2 Studies

The thesis evaluates the research question with two studies, one looking at this with greater breadth over a shorter period (Study 1) and a second with greater longitudinal depth over a longer period of use (Study 2). To be consistent with a self-directed learning setting, system use by participants was entirely voluntary and supplementary to ongoing courses within the School of Electronic, Electrical and Systems Engineering (EESE) at the University of Birmingham (UoB). A limitation imposed by the school was that access to enough participants to investigate all aspects longitudinally would not be possible at the time of evaluation, due to numbers enrolled and competition for other initiatives. However, access to a subset of students in one module was granted and this is reported as Study 2. It is important to retain evaluation relating to cross domain usage in addition to contrasting between the cases where students were actively engaged in learning the domain content at the time of use, versus it being more indirectly applied to their current academic work. The studies should also cover a period where the model has sufficient information in it for there to be navigational decision points when engaging with the model's content, and therefore investigation with a larger sample over a shorter period was still sought (Study 1).

5.3 Beyond the Scope of Evaluation

It is worthy to note issues that are beyond the scope of this research, such as:

(i) Any potential learning gains. An open learner model alone is not designed to teach, but to support students alongside their learning – factors to regulate the student's learning environment are not intended to be controlled, as this would likely affect what is under measurement as part of the research question.

(ii) A comparison of the effectiveness of the open learner model in drill down form, as compared to the same informational provision with other interface techniques that make all information available concurrently. Literature suggests that for reasons of cognitive load and informational brevity that drill down approaches are part of a good and pseudo-established open learner model design practice. Overloading the student with information in a flat structure or implementing only a coarse-grained interface goes against what is suggested in the literature as best for promoting metacognitive practice. The evaluation in this thesis therefore looks more at context of use.

(iii) A comparison of the effectiveness of the open learner model in drill down form, as compared to an open learner model that is opened at an overview level only, or nonprovision of a technology-based solution. Literature suggests that while even simple open learner model visualisations can be effective for learners, removing the opportunity to access the detailed information, when there is an informational need when problems are apparent, changes the formative assessment opportunities afforded. Controlling this factor would change the evaluation to be about information provision. It would also omit content relevant for mirroring students' cognitive tasks at different levels of detail, which literature suggests is important with respect to metacognition.

5.4 Participants

The research question requires an investigation consistent with other OLM evaluations, for comparison. Many of the case studies presented in Chapter 4 are at university level and a participant group of university students is consistent with the preferences of the University of Birmingham ethics policy (also see Section 5.7). As Chapter 3 suggests, OLMs are particularly suited to STEM disciplines, and so contact was made with the School of Electrical, Electronic and Systems Engineering at the University of Birmingham for the purposes of evaluation. With agreement of the school, participants were sourced from within programmes of study that the school provides, with the suggestion that the domain content of the OLM should be drawn from the commonalities of all tuition streams and so all should have familiarity and direct benefit from the technology. A requirement was introduced by the school, consistent with the educational benefits detailed in the literature, that access to the students would be granted if the technology containing the OLM was able to generate its own questions, and this is incorporated into the design (see Chapter 6). The school's lecturers provided a wealth of support and expertise to ensure that the domain content of the OLM system developed (Chapter 6) was accurate, appropriate and timely for their students' tuition. As such, the technology developed is bespoke to students of the school of EESE. The participants available for evaluation are drawn from volunteers within this school only and in modules where the lecturers supported the initiative.

The school has approximately 250 undergraduate students across 4 year groups, and students undertake a range of degrees from Electronics and Electrical Engineering through to Computer Systems Engineering. This includes BEng, MEng and MSc programmes of study, with additional direct entry students in Years 2 and 3. Most students are in the age range 18-25. A foundation for all programmes is participation in modules of basic engineering mathematics (algebra, matrices, integration and differentiation etc.) and Boolean logic (logical operations such as AND, OR and NOT and algebraic manipulations of these operations). These building blocks form a foundation for other modules including power electronics, computer programming in a range of languages, group engineering projects, robotics, datamining, artificial intelligence and educational technology. As covered in more detail in Section 5.8 no additional details are recorded about the personal profiles of student participants within the school, including the gender balance, demographic or academic performance.

5.4.1 Participants for Study 1

A total of 60 volunteer third year undergraduate students studying Computer Systems Engineering at the University of Birmingham participated in the 1 hour study. Students were enrolled on a course about educational technology and had covered basic engineering mathematics and Boolean algebra earlier in other aspects of their degree programme, and so had a level of prerequisite knowledge. They were using the domain content in other modules, although this was not directly part of the curriculum in the module in which initial contact was made with the students. Participants are excluded from the sample who did not stay for the full duration of the lab session or who did not give consent for use of their data. Students had previously used OLMlets (Bull and Mabbott, 2004) alongside other courses, and so had some level of familiarity with working with an OLM. Participants were split between the four sample groups: MA(U):16, MA(I): 16, BO(U): 13, BO(I): 15, as described in the methods section of Section 5.6.1 .

5.4.2 Participants for Study 2

A total of 27 volunteer first year undergraduate students from the school of EESE at UoB participated in an 8 week evaluation, alongside their main studies. Participants were all actively studying the basic engineering mathematics module which is common to first year degree streams, whether participants subsequently specialise in electronic/electrical engineering or computer systems engineering. The domain content of the system exactly mirrored their course structure (including working with expressions, logarithms, matrices, basic calculus and probability). Contact with the students about use of the technology was made through the course lecturer at the start of term. Subsequently, one hour lab sessions were available each week during which participants had access to the researcher so that they could clarify anything they needed to regarding the OLM technology. As per Study 1, participants are excluded from the sample who used the system on fewer than three occasions across the term, or who did not give consent for the use of their data. Students were using OLMlets (Bull and Mabbott, 2004) alongside other courses, and so had some other exposure to OLM technology at the same time.

5.5 Materials

Both Study 1 and Study 2 made use of the same materials for evaluation. These included: an OLM system with which students may interact, and which recorded metadata about student interaction (Section 5.5.1); a post usage survey to capture participant perceptions about their use of the OLM (Section 5.5.2); and an interaction log analysis suite to aggregate and visualise learners' interactions for the purpose of analysis (Section 5.5.3).

5.5.1 Technology to Record Data

The survey of technologies (Chapter 4) showed existing examples of OLMs implementing a drill down approach. For the research considered in this thesis control is required over what information is presented at each level of drill down and the deepest level is required to show specific beliefs inferred from students' behavioural models (consistent with the granularity taxonomy presented in Table 8). This creates a need for one of the systems in Chapter 4 to be modified or a new system to be constructed that is tailored specifically to the needs of the research. The latter approach is adopted, and a tailored open learner model is coded from scratch. To evaluate the use of a drill down approach, the technology and analysis must be able to have clean data indicating access of information at each level, and so an implementation of the OLM is achieved whereby the method of data collection does not disrupt or bias student use of the OLM, and where each navigational action has a distinct unambiguous meaning. The design for an open learner model implementing a drill down approach is included in Chapter 6 and its data falls into two categories:

(i) that which is required to present the learner with an open learner model – this data is specifically about students' current understanding and is the behavioural model.

(ii) that which is required for analysis but is not directly presented to the learner – this comprises the diagnostic model (required to calculate self-assessment accuracy) and a log of interactions with the system (required to give a description of interaction behaviour).

The OLM implements a drill down approach in its interface design, consistent with the requirements of the research question detailed in Section 5.1 (also see the design in Section 5.6.4). The navigational transitions between elements of the interface are used to generate metadata to indicate transitions between different depths of drill down or states of interaction. Additionally, these also record a timestamped state of the student's behavioural and diagnostic models at the point of access. This metadata feeds the analysis tool described in Section 5.5.3 and analysis methods of Section 5.6.5 .

Automatic question generation is implemented to counteract memorisation of answers, this is also coded from scratch with academic content coming directly from the participant students' university modules. Students' responses to multiple choice questions is the method from which the learner model is updated. The action of submitting the question response is combined with learner self-assessment to indicate whether students believe their response is correct. This approach is also taken by (Bull and Pain, 1995; Kerly et al., 2008). The intervention, as such, keeps a behavioural model (students applying domain knowledge, to demonstrate understanding) and a diagnostic model (students self-assessing their abilities in completing the task) both to the same precision, and it is updated through the same implicit action. A full discussion of these issues is given in Chapter 6.

5.5.2 Survey to capture participant response

To understand student usage, utility and perceptions of the system a post usage survey is required. The survey is post usage as not to influence student interaction and was distributed on paper at the point at which learners indicated that they no longer wished to make use of the intervention method at the end of the evaluation session(s). Use of a 5-point Likert scale is common in the field of open learner modelling for such evaluative purposes (e.g. Ahmad and Bull, 2008), and this is sometimes reduced to a 3-point scale in analysis. Students are asked to indicate their agreement with statements on a scale of 5 (strongly agree) to 1 (strongly disagree). Where questions are about visualisations or information at different levels of granularity, the same questions are asked for each instance, such that comparison may be drawn. The full survey is included in Appendix 3, and it covers:

- Student acceptance of each of the visualisation components at each level of granularity and also acceptance of the drill down structure of the OLM. *Contributing to Q1a.*
- Student perceptions of their self-assessment abilities. *Contributing to Q2a, Q2b.*
- Informational interests in OLM content at each level of granularity, contributing to students' interest in their problems and problem solving attempts. *Contributing to Q3a and Q3b.*
- Student reasons for viewing the OLM in the context of updating it, to give indication of potential evidence of student focus. *Contributing to Q3c.*

5.5.3 Technology for automated analysis

The technology was designed and implemented to automatically capture interaction data (Chapter 6). The research question requires some pre-processing and analysis of the raw interaction logs, to provide information for further analysis in both a statistical and visual form. As part of the technical development work that has taken place for this thesis, algorithms have been implemented to calculate the state of the student behavioural and diagnostic models throughout interaction, differences between behavioural and diagnostic models at each point, levels of completeness of the models, detection of interaction state (as is described in Section 5.6.4) and transitions between the different interaction states. This is done to determine the significance of student interactions at each point in time given all possible other actions and their informational states. Algorithms written aggregate this information and are used provide the data for the analysis presented in the graphs in Chapter 7 and Chapter 8. Additionally, a front-end visualisation tool is connected directly to this and screen shots from this are included in the analysis[1]. All other figures are drawn with the data output of these analysis algorithms only. Statistical significance tests are completed separately to the automated analysis.

5.6 Methods

As described in Section 5.4 the evaluation is split across two studies. This section describes: the methods used during the capture of data as part of the evaluation (Section 5.6.1 and Section 5.6.2); the underlying concept behind how students should be free to interact with the technology that will lead to interaction data where students' actions are distinct at each level of drill down (Section 5.6.4); and the analysis methods/analytics required to answer the research question (Section 5.6.5 to Section 5.6.7).

5.6.1 Study 1: "Broader Perspective"

Students were first introduced to the technology by means of a small group demonstration (3-5 participants at a time) on a test account by the investigator. The bullet points for this introduction are included as Appendix 4 and were adhered to consistently on each occasion.

[1] Figure 10, Figure 11, Figure 12, Figure 13, Figure 20, Figure 21, Figure 22, Figure 23, Figure 32, Figure 33, Figure 34, Figure 35 and Figure 36 are draw using this software.

The introduction included a mechanics of how to operate the technology and how to interpret the information presented in the OLM. Care was taken to ensure that bias was not given during demonstration, for example towards selecting items for drill down that were only problematic. The technology was presented in a factual way, indicating that it was for formative assessment and had domain content that related to specified modules within their degree programme. No details of educational benefits of interacting with this technology were disclosed, no statements were made about whether the technology would be useful in their learning, and no prescribed uses for the technology were suggested. Participants were allocated to each of the four experimental groups at random, were free to ask for clarification on how to use the system at any point, and were informed of their right to non-participation for research purposes and their right to withdrawal (see Section 5.8 for an extended discussion).

Students independently made use of the system described in Chapter 6 for the period of one hour. At the end of the one hour session students were asked to complete a survey capturing their perceptions of use (Appendix 3), completed a consent form (Appendix 1) and were given debrief information indicating how their data would be used and the purpose of the study (Appendix 2).

With a view to presenting some qualitative context in terms of volumes of information and elaborating on the potential in terms of a non-domain specific finding, the system was used in two different STEM domain areas (basic engineering mathematics; Boolean logic). Following on from the open learner model literature of Chapter 3, initially, an OLM would be empty, until such a time as there is relevant information to show to the learner. It may also take a period of interaction until there is sufficient information in a learner model before it is at full accuracy. In a self-directed learning context, where students are free to choose which part of the model to interact with and update, it may be the case that the model will not be complete, and the same part of the model may not be returned to over time. For a consideration of impact on self-assessment accuracy and tasks regulatory to cognition over time this requires multiple data points. It is not the desire to restrict student interaction at any point, however to overcome this potential limitation, use was made of two experimental groups: one with a complete model; and one with an initially empty model.

Table 6: experimental groups.

	Basic Engineering Mathematics	Basic Boolean Logic
Students build the model during interaction, and may reference it at any point (uninstantiated)	MA (U)	BO (U)
Students' models are already built by the first point of interaction. Students may continue to subsequently provide information to update the model. (instantiated)	MA (I)	BO (I)

This leads to a 2 x 2 design for experimental groups (2 domains x 2 contexts of use), as shown in Table 6. It is not the intention of the thesis to quantitatively compare these, as this is not required to answer the research question, however some qualitative comparison will elaborate the context of use and consider which aspects may generalise.

In terms of the implementation of these there are two groups that use the system in the domain of basic engineering mathematics (code "MA"), with a domain of 3 topics and 9 concepts, and two groups that use this in the related engineering domain of basic Boolean logic (code "BO"), with a domain of equal size, 3 topics and 9 concepts. Likewise, in considering the effect of a large amount of updates to the model, two of the groups (with code "U" - uninstantiated) are permitted to use the model at any point during interaction, using the accretion principle (see Kay, 1994). The two remaining groups (with code "I" - instantiated) are required to provide enough information for the model to be mostly complete at the point of first access, thus all information is new, simulating an import of data.

5.6.2 Study 2: “Longitudinal Perspective”

As per Study 1, students were first introduced to the technology by means of a small group demonstration (3-5 participants at a time) on a test account by the investigator. The bullet points for this introduction are included as Appendix 4 and were adhered to consistently on each occasion. The introduction included the mechanics of how to operate the technology and how to interpret the information presented in the OLM. Care was taken to ensure that bias was not given during demonstration, for example towards selecting items for drill down that were only problematic. The technology was presented in a factual way, indicating that is was for formative assessment and had domain content that related directly to their current module of study. No details of educational benefits of interacting with this technology were disclosed, no statements were made about whether the technology would be useful in learning, and no prescribed uses for the technology were suggested. Participants were free to ask for clarification on how to use the system at any point and were informed of their right to non-participation for research purposes and their right to withdrawal (see Section 5.8 for an extended discussion).

Students independently made use of the system described in Chapter 6 at times of their choosing over an 8 week period. One hour of contact time with the investigator per week was made available to the entire group, and all use was entirely optional. At the end of Week 8 students were asked to complete a survey capturing their perceptions of use (Appendix 3) (24 of 27 participants responded), completed a consent form (Appendix 1) (27 participants responded) and were given debrief information indicating how their data would be used and the purpose of the study (Appendix 2).

5.6.3 General Evaluation Methods and Terminology

In addition to the survey instrumentation (Section 5.5.2) throughout the analysis use is made of statistical methods and specific terminology for referring to the different depths of drill down (elaborated in Section 5.6.4). For inferential statistics, three main types of test are used: (i) t-tests for before and after conditions of matched pairs, (ii) Spearman rank correlations for comparing matched pairs in two different data fields, (iii) ANOVA testing where comparison is required between multiple fields. Inferential statistics are used to a lesser extent for validating user acceptance and are exercised to a greater extent in the analysis of impact on self-assessment and regulation of cognition. Tests are non-parametric unless otherwise stated and are reported according to their significance level (5%, 1%, 0.1% etc).

Table 7: analysis terms for OLM inspection and their prerequisite navigational paths.

Term	**Description**	**Prerequisite Navigation**	**Informational Type**
Overall	A summary of the topics, on first navigation to the model. Prerequisite navigation: none	Cycle of answering multiple choice questions	Domain structure representation. Collapsible tree. Model shown through coloured nodes.
Topic inspection	Breakdown of a topic to show the model for constituent competencies	Overall inspection	Domain structure representation. Collapsible tree. Model shown through coloured nodes.
Concept inspection	Breakdown of a concept to show the weighted model	Topic inspection	Weighted model representation for a specific concept. Skill meter representation.
Belief inspection	Breakdown of a concept to show the inferred beliefs, as per the weighted model	Concept inspection	Specific article of information in the learner model. Mathematics notation and descriptive text.
Domain inspection	Breakdown of a concept to show the domain content (system beliefs); beliefs that the learner model would show if the questions were answered correctly	Concept inspection	Specific article of information in the domain model. Mathematics notation and descriptive text.

The arrangement of the software (see Section 5.6.4 and Chapter 6) allows learners to drill down into their OLM, starting from a high level overview of the topics, through to detailed beliefs inferred from multiple choice questions answered. From one level it is possible to move to the next more detailed level. In the analysis the following terms used to refer to each granular level are as per Table 7.

5.6.4 Drill Down Method Design

Table 3 in Section 4.1 summarises the different granularities to which a learner model may be opened to the learner, using 7 different levels. For an implementation of an OLM that is tailored to the research question this thesis proposes that all these different levels of granularity do not need to be present in order to proceed with the investigation, but a sufficient number of levels need to exist in order for any potential baseline behaviour associated with each to be clearly identified. Reducing the number of levels reduces the complexity of the analysis and of the required implementation, and it also should be considered that an investigation into epistemology (as the finest level of granularity) is an extension item for the research. The thesis therefore proposes that this be simplified to the levels of granularity as indicated in Table 8.

Table 8: granularity of learner model externalisation.

Coarse | | | | | V Fine	• Level related information; i.e. ability to correctly apply a specific rule; whether something is known; aggregation of other framework items, or partial model. • Level related information broken down into sub-nodes (e.g. as per a domain structure) • Level related information justified in terms of other model information such as specific model weightings • Examples of applied understanding, specific learner beliefs and inferences. E.g. specific rule applied.

To permit information of each type to be visualised at interface level requires the content of the learner model to be at the finest grain size in Table 8. In the technology in this thesis this is specific articles of knowledge and applied knowledge – i.e. specific beliefs of which the learner has been able to apply.

This implies that the interface should also mirror this granularity in its navigational structure and implementation of drill down functionality. Viewing specific articles of information could be seen as states during interaction (Figure 2). It is important to note that the learner is free to provide further information at any point and to view any aspect of the information that they wish. All navigational decisions are taken by the learner and are in part used to describe the context of use. The learner remains in control (see best practice literature on self-directed learning, Section 2.4 , and open learner modelling, Section 3.3). In Figure 2 the finer grain sizes of information are to the right of the diagram, and the coarser grain sizes to the centre-left. All navigational possibilities are shown by a directional arrow. Access to the domain content at the level of the greatest detail stems from the requirement for the access of contextual information (see also Section 3.6).

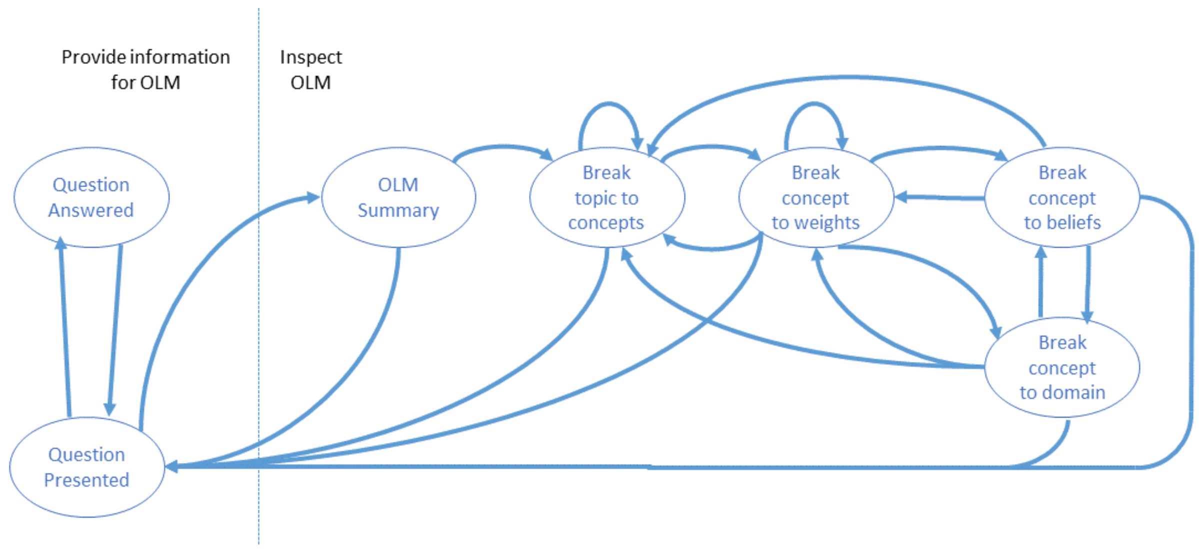


Figure 2: interaction state model.

As a method for describing the deepest level to which the model may be drilled upon any given single access of the resource, the workflow of Figure 3 can subsequently be used to describe this, providing contextual information for the use of drill down, for analysis.

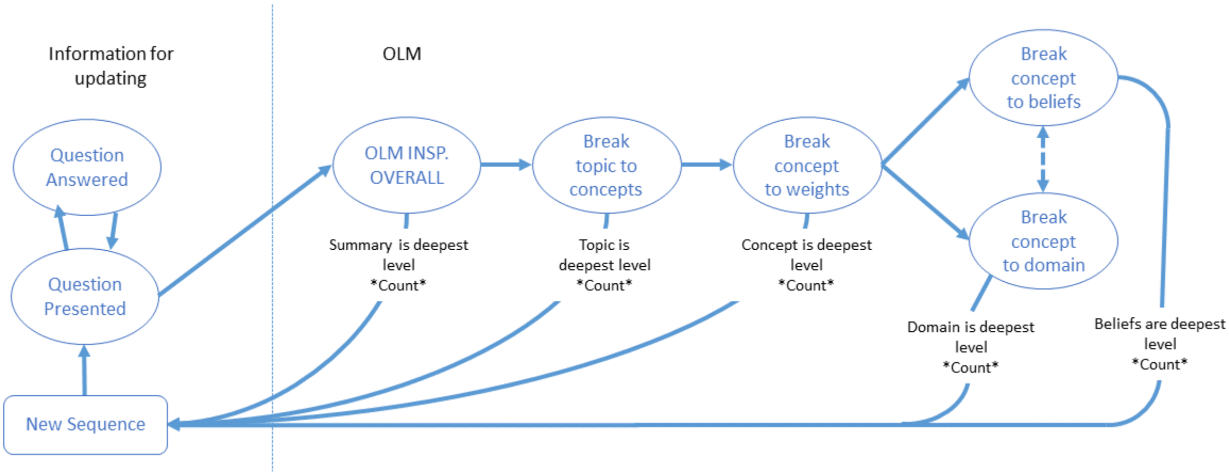


Figure 3: workflow for identifying the deepest level of OLM inspection, per access.

5.6.5 Evaluation methods for Q1: how is an open learner model with a drill down approach used?

Chapter 4 has suggested that it is appropriate to ask how an open learner model with a drill down approach is used as: this is a gap in the literature; OLMs are open to the learner for metacognitive benefit and often implement a drill down structure; and it is valid to confirm conditions under which metacognitive benefit may be achieved in an OLM. To do this a behavioural description of use over time is required and it is standard in OLM research to first consider students' acceptance of the technology to contextualise this. This thesis asks:

5.6.5.1 Q1(a) Is an open learner model with a drill down approach accepted by its users?

To address Requirement 8 and contextualise much of the analysis an element of user acceptance is considered for those who chose to participate in the study.[2] Within the field of OLM it is standard to look at the extent of usage and verify whether elements are understood and useful, as part of user acceptance. The analysis therefore comprises a behavioural description of student use and quantitative description of their perceptions of use. This includes a behavioural description of use that is relevant to also looking for evidence of support for problem solving as a task regulatory to cognition (cross over with Section 5.6.7) by considering the numbers of questions answered by students in succession/frequency of use of the model. In order to consider the extent and shape of usage:

- Frequency of use statistics will be presented for the different features (different levels of drill down) in the OLM. This should be contextualised by the number of questions answered, giving an indication of how much information the model is based on, and also a basic pattern of life of how it is updated. Standard deviations should be used to show variance.[3] This contributes a picture of the extent of use of each element of the OLM to contribute to generalisable patterns in descriptions of use.
- Survey content (Appendix 3) capturing students' perceptions on whether aspects of the drill down approach are understood and useful, and additionally students' rationale for accessing the information. This contributes towards answering whether there is student acceptance.

---

[2] It may also be considered that the remainder of the population of the school did not accept the technology, but it is not known whether this might be because of not being informed of the opportunity to participate, the initiative not being supported by course lecturers, because of competition for participation in other initiatives, because of a rejection of the idea of the use of a tool of this type for formative assessment purposes, or for other reasons. These members of the population would also not be informed of the precise capabilities of the technology, and so user acceptance is considered from the perspective of those who did choose to participate.
[3] No statistical significances are intended to be used for this measure, as acceptance is more subjective, and no benchmark statistical definition exists for what user acceptance of an open learner model might be.

5.6.5.2 Q1(b) Is use of the drill down approach in the open learner model consistent across time?

To further consider user acceptance and establish any baseline interaction behaviour, usage should be considered over time. To describe any differences between how coarse and fined grained parts of drill down are used, this description of use should compare these elements individually. This will also help give indication of the temporal dimension of conditional use of drill down. Appropriate analytics should include:

- Cumulative distribution frequencies showing the extent of use across time within the sample. To allow students to be compared directly and aggregated. These should be normalised (thus the time of the first interaction is 0.0 and the final interaction 1.0; all periods of inactivity greater than 30 minutes are removed, as not to distort the data through the inclusion of dormant periods between usage episodes). This contributes a temporal description of the consistency of use of the OLM and of drill down across time. It allows the consistency of interaction at different levels of granularity to be compared. In addition to contributing to the baseline description of use, it also contributes to Q3 (Section 5.6.7).

5.6.5.3 Q1(c). Is drill down always used when inspecting the open learner model?

To further consider how the OLM and drill down are used, and to help define the baseline student interaction with the drill down capability, a consideration should be given as to whether this is always used or whether this is conditional. Analysis comprises:

- An evaluation of the deepest level to which the model is drilled, per inspection, to give an indication of any conditional use of the deeper levels of drill down.
- An evaluation of participants' interactions in terms of the interaction state model that underpins the system design, as shown in Section 5.6.4 . This contributes a formal description of what students did, in the context of all navigational possibilities, and in the context of the intended underlying structure of the system that is constructed to elicit distinct actions at each level of drill down. This also further contributes to Section 5.6.7 .

5.6.6 Evaluation methods for Q2: what is the impact of an open learner model with a drill down approach on self-assessment accuracy?

Chapter 2 highlights that self-assessment is a regulatory process that contributes to metacognitive development, and Chapter 3 indicates that this is one of the main reasons for opening the learner model to students. Chapter 4 has also suggested that it is appropriate to ask how an OLM with a drill down approach impacts on aspects of self-assessment, particularly in terms of self-assessment accuracy, as: this is an ongoing line of enquiry; previous studies have shown improvements in student self-assessment during OLM interaction; and that it is valid to investigate ways that this can best be supported. To understand the impact on self-assessment it is appropriate to consider how this changes across time, and to understand the specific role of the drill down structure. It is appropriate to consider how this varies with the extent of its use. The analysis comprises:

5.6.6.1 Q2(a). Will student self-assessment accuracy increase over time when using an open learner model with a drill down approach?

Consideration of how accuracy of self-assessment changes across the period of exposure to the OLM technology requires appropriate measurement points to be identified. As the model is initially empty for some experimental groups, self-assessment accuracy may not be calculated at the very start of interaction. Literature has suggested that a comparison of students' behavioural and diagnostic models to the same level of precision is an appropriate method by which to calculate an accuracy measure. As per Section 5.6.5.2 this also requires interaction to be normalised, where $t=0.0$ is the first interaction and $t=1.0$ is the final interaction. This is further considered in Appendix 5. The analysis therefore reports:

- Change in student self-assessment accuracy between an appropriate first measurement point and the final interaction time. A two tailed t-test may be used to determine if changes are significant. This contributes a direct answer to whether self-assessment accuracy is changed.

The analysis here potentially also needs an additional piece of context to interpret it. There is no assumption that all learners will have the same level of competency in self-assessment, and no assumption that all students will have made equal use of the OLM in general. A full evaluation of existing self-assessment ability of participants is not possible due to externally imposed conditions; however, a proxy measure is considered in terms of participants indicating their self-assessment ability as a survey item. The analysis therefore also additionally considers the following as context:

- Can any change self-assessment ability be represented by the proxy measure of students indicating their self-assessment ability? A Spearman rank correlation should be used to determine the significance of this. This will identify if student stated self-assessment ability can be used as a proxy measure to identify those who will be impacted to a greater extent through OLM use.
- Is any change in student self-assessment ability related to the frequency of use of the OLM? A Spearman rank correlation should be used to determine any significance. This contributes to a consideration whether impact on self-assessment is affected by the baseline pattern of frequency of use, as identified in Section 5.6.5 .

5.6.6.2 Q2(b). Will student self-assessment accuracy increase to a greater extent when greater use is made of drill down?

Drill down is something that is optional for students to engage with, and literature highlights that its use is conditional. Building on the conditional use analysis considered in Section 5.6.5

consideration should be given to how the extent of use of the drill down process impacts on self-assessment accuracy. The analysis should include:

- The extent of use of each level of drill down cross correlated with any change in self-assessment accuracy. A Spearman rank correlation should be used to determine the significance of this. This contributes to whether the use of drill down is related to changes in self-assessment accuracy.

5.6.7 Evaluation methods for Q3: what is the impact of an open learner model with a drill down approach on the support for the regulation of cognition?

Chapter 3 indicates that metacognitive benefit is fundamental to the argument of why learner models are opened to students, and Chapter 2 indicates that metacognition comprises knowledge of cognition and the regulation of cognition. It further highlights that tasks important in the regulation of cognition have measurable elements which may be used to determine the presence of behaviour that is metacognitive in nature, such as problem solving or comprehension monitoring. The thesis therefore looks for consistency with these behavioural elements to determine if an open learner model implementing a drill down approach has impact on the support for the regulation of cognition. As stated in Section 2.1 this includes evidence of recognising or identifying problems, detailed inspection of problems, application of strategies such as trial and error or hypothesis testing and monitoring of feedback to determine the success of actions. To complete an analysis of the presence of these elements, the thesis builds on the behavioural baseline activity analysed in Section 5.6.5 and additionally considers whether students drill down into the model to inspect information about their problems, whether this is distinct from other factors such as their perceptions of their ability to correctly apply domain knowledge, and whether there is evidence of focussed interaction with one area of a domain at a time, consistent with the strategy and feedback phases of problem solving. The analysis includes:

5.6.7.1 Q3(a). Will students use drill down in the OLM to inspect information about problems?

This section of the analysis looks for supporting evidence for students being able to identify problems, and to perform careful inspection/observation of their problems as part of the problem identification and problem structuring phases of problem solving. The measures are:

- The difference between the (behavioural) learner model state for an item inspected and the average state of the learner model, at each point of access. This will determine if the item inspected is problematic as compared to other navigational possibilities. This should be done across each level of drill down to indicate whether students are drilling into the model to identify problems and whether there is evidence of detailed inspection of the model where problems are greatest. The learner (behavioural) model state ranges from 0.0 (not understood) to 1.0 (understood), on a per concept basis, using the modelling process described in Section 6.3 . The nodes at 'overall' level are aggregated, thus partial models. At any one point in time the model will have an average state, and the part of the model being drilled down into will also have a model value. This thesis considers the discrepancies between the two, to identify whether students are drilling down into items that are problematic, in the context of their model, and the magnitude of the difference of these. Calculations are first done on an event-by-event basis per participant, and then averaged over participants - the SD values thus refer to the variance between participants. A two tailed t-test is used to determine the significance of the action. Significances between the levels of drill down (drilling deeper with greater problems) may be calculated using a one way ANOVA (related).
- To identify whether this is consistent with students' perceptions of usage (i.e. whether they're consciously inspecting their problems) survey items should be included to identify students'

perceptions of their reasons for using the OLM, in the context of the reasons for which the model is initially opened to the learner (see Appendix 3). The same questions should be asked about each level of drill down to determine if there are any differences of use at each level of granularity. This contributes confirmation of whether students are consciously using the OLM to identify problems.

5.6.7.2 Q3(b). Will students use drill down in the OLM to inspect information about areas of uncertainty?

As part of a consideration of whether students are using the OLM to inspect content about problems is identify whether each inspection relates to areas that students perceive as problematic before accessing the open learner model (i.e. before being presented with confirmation that something is a problem). In the design of the system this translates into a comparison of how the diagnostic model of student self-assessments relates to students' navigational actions and their use of drill down. Similar to the analysis of Section 5.6.7.1 this includes:

- The difference between the (diagnostic) learner model state for an item inspected and the average state of the (diagnostic) learner model. This will determine if the item inspected is problematic as compared to student perception of their ability to correctly apply domain knowledge. This should be done across each level of drill down to indicate whether this is a motivating factor for the use of drill down in the OLM. As per the previous section, the learner (diagnostic) model state ranges from 0.0 (not understood) to 1.0 (understood), on a per concept basis, using the modelling process described in Section 6.3 . The nodes at 'overall' level are aggregated, thus partial models. Any one point in time the model will have an average state, and the part of the model being drilled down into will also have a model value. This thesis considers the discrepancies between the two. Calculations are first done on an event-by-event basis per participant, and then averaged over participants - the SD values thus refer to the variance between participants and a two tailed t-test is used to determine the significance of the action. Significances between the levels of drill down (drilling deeper with greater problems) may be calculated used a one way ANOVA (related).

5.6.7.3 Q3(c). Will students use an OLM with a drill down approach to focus on one domain area at a time?

As part of an analysis of whether any baseline student interaction is consistent with problem solving strategies, calculations need to be made to determine whether information is being inspected in the OLM immediately after its update as all navigational actions are under the learner's control. This should be synthesised with the pattern of use identified in the analysis of Section 5.6.5 and be supplemented with student perceptions of their interest in different aspects of information when inspecting the OLM. The analysis should include:

- Identification of how recently an item was updated when it is selected for inspection in the OLM. This may be calculated by considering the number of question blocks ago the domain area was encountered and should be calculated for each level of drill down to determine any differences of use at each level of granularity. Means and medians should be used to identify any bias in the dataset, and SD values should be used to indicate consistency. The significance of the pattern across the different levels of drill down (i.e. when drill down is used is the information more/less recent) should be indicated by a one way ANOVA (related). This contributes to whether there is any indication in the baseline interaction of students working in a focussed way and whether the OLM is being used for immediate feedback.
- Survey content about students' perceptions of why they access information in the learner model, in the context of updating it. This should be used to indicate whether students

consciously desired information about the last domain area updated at the point at which the model is inspected.

5.7 Requirements

The literature review identified 8 requirements that had implications for the evaluation of the research question and the design of the OLM technology with which learners interact. These have been addressed in the design of the evaluation (this chapter) and the OLM technology (Chapter 6) as follows (Table 9):

Table 9: actioning of requirements.

ID	Description/Rationale	Action
R1	Tasks such as evaluation, comprehension monitoring and problem solving or debugging relate to the regulation of cognition, as a subset of metacognition that is more measurable. The technology should allow for evaluation aspects of these.	The OLM technology is described in Chapter 6. It logs all student interactions and implements a drill down approach in the OLM, where information is split according to granularity. The OLM identifies the state of student understanding as compared to the domain. Students may answer questions at any point to update the model and all navigational decisions are under the control of the learner. This approach allows for behaviour and interest in specific elements of information to be recorded, giving a detailed description of student interaction behaviour, including interest in specific problems.
R2	Self-assessment is a regulatory practice that contributes to metacognitive development. Consideration should be given for how this changes across time to understand its impact.	The OLM keeps a models of student behaviour and self-assessment to the same precision across time. Appendix 5 identifies a method for determining suitable measurement points as the model is initially empty and calculation of learner model accuracy across time.
R3	An appropriate point for OLM use is during attempts at self-directed learning, this is a point during which impact on self-assessment or aspects of metacognition might be expected in technology such as this.	The deployment setting for the OLM is for 8 weeks alongside a university course and during shorter lab sessions at university level. Students use the technology under their own direction and the curriculum content is taken directly from their university course.
R4	Different visual/informational forms are needed, tailored to specific (relevant) cognitive tasks. Graphical and textual forms are most common, and this allows the OLM to be consistent with existing implementations.	The design of the OLM (Chapter 6) caters for overview metrics using coloured/labelled nodes, through to skill meter representations showing the extent of understanding, through to specific beliefs in both textual and domain specific notation. These are consistent with other OLM examples and provide external representations of the learner model that would be appropriate at each level of granularity.
R5	Use of a drill down approach allows for supporting tasks and capturing information relating to the regulation of cognition, and also a way for learners to access contextual information required to	The OLM technology (Chapter 6) implements a suitable drill down approach, consistent with examples in the literature. The most detailed view (beliefs) also presents a representation of the domain in the same form for comparison by the learner. The OLM technology caters for

	interpret visualisations. An evaluation of the consistency of use of drill down will be required to describe how it applies to OLMs.	capturing of suitable information for tasks relating to the regulation of cognition (see R1) in forms that are appropriate for smaller cognitive tasks (see R4). The consistency of use is addressed as part of the first sub-research question.
R6	Information should be split broadly according to cognitive task/granularity and transition between these should be supported by an active navigational approach. This will describe how drill down is used.	The OLM presents views of information suitable for tasks such as overviewing, showing the extent of understanding and detailed beliefs (see also R4). The OLM interface is designed such that mouse clicks indicate a desire to inspect information in each category and interaction logs are used to quantify learners' behaviour.
R7	The system must maintain and record comparable models of behavioural (e.g. performance) and diagnostic (e.g. self-reporting) data. This will allow evaluation of changes in aspects such as self-assessment accuracy and a temporal description of how learners are approaching the domain content.	The OLM technology (Chapter 6) keeps a behavioural model (maintained by students answering questions on a chosen area of the domain, and their response being compared to the domain model) and a diagnostic model (maintained by student self-assessment of their ability to answer a question correctly, at the point it is submitted to the model). These are to the same precision, and all actions are timestamped, such that the model of a student may be reconstructed during analysis to the state it would have been at any point during interaction.
R8	User acceptance of the open learner model and drill down is required to contextualise any usage behaviour, it cannot be assumed.	An evaluation of user acceptance is part of the first sub-research question. It concentrates on whether the OLM and the drill down approach are used, understood and useful. Subsequent areas of analysis are informed by this validity.

5.8 Ethics

This thesis next considers aspects of ethics covering the: (i) automated collection and storage of data for research; (ii) data recorded about students' personal profiles; (iii) conduct of the evaluation; and (iv) conduct of the analysis. The outlined research was approved by the University of Birmingham Ethics Committee on the 1/12/2009 with reference ERN_09-243.

5.8.1 Automated Collection, Storage and Access to Data

The OLM implemented for research is as a web-based resource that is accessible from anywhere and therefore data security and the protection of learners' data is an important issue. Each user account is password protected, contains only the information of that one user (there are no collaborative aspects or information sharing between learners as part of the investigation) and is accessed only via the learner's pseudonym (their chosen username), rather than name. The data stored in the back end is on secured servers and all data is stored by using an anonymised unique identifier (random number) as a key, and all passwords and emails are stored as MD5 hashes and so are not human readable. The right to privacy is addressed as the user and their personal information are not identifiable in the data stored, and only the researcher and ICT administrative staff have privilege to access the server. The right to equal access is addressed as students retained electronic access to the technology for the remainder of the academic year after the study had completed, whether they participated in the study or not.

The research required metadata to be generated about student use. This included each mouse click, time durations to complete actions and summaries of information present on the interface, in addition to information explicitly given to the system (e.g. question responses, self-assessments) and information inferred about the learner that they had access to via the OLM interface (e.g. specific beliefs). Learners were informed of this data collection during the introduction to the technology and it was possible to view the raw data collected by request after the study. Only data intended to be used in this research study was captured, and it is only accessed for research purposes where written consent is obtained. The data is to only be stored for the length required by the documentation of this study – this is specified as 5 years after the research has been completed, including any publication arising from the results.

Student survey data is captured in hard copy and is stored anonymously in a secure location, accessible only by the researcher/supervisor. Cross referencing the sample number with the consent form is the only way to link to the origin of the data.

The consent forms (Appendix 1) completed by users are the only items bearing the student's identity and would only be accessed to obtain the recorded sample number should the student subsequently wish to withdraw from the study and the removal of their data from the dataset is required. Consent forms have been stored securely by the supervisor throughout the duration of the research.

5.8.2 Access to Data about Students' Personal Profiles

All participants were adults of university age and able give informed consent for the use of their data without a third party. This is consistent with the University of Birmingham ethics policy for PhD studies which suggests avoiding working with participants under the age of 18 unless this is essential to the research's line of enquiry.

The information collected about participants is solely their anonymised interaction data through usage of the OLM and their responses to the post-usage survey (Appendix 3). The participant's name is collected on the consent form as to be able to identify the participant with a sample number should they wish to withdraw from the study at a later point in time. Subsequent to discussion with course lecturers and the University Ethics Committee, no other personal information is recorded, and within the research there is no cross correlation between students' interactions and other areas of their personal, demographic or academic profile. This means that it is not possible to consider issues relating to gender, academic performance, wider areas of study or potential learning difficulties, within the analysis.

5.8.3 Conduct of the Evaluation with Participants

The participants were introduced to the system and its concepts and had opportunity to ask any clarifying questions about its purpose, mode of operation and data it was capturing prior to use. The students were under no obligation to participate in the study and were free to use the OLM throughout the academic year at any point even if they were not taking part in the study. Student participation was extra to the commitments of their courses and carried no summative weighting. At the end of the sessions participants gave written consent for their data to be used for the purposes of research (Appendix 1). Consent was obtained at the end of the study so that students were more informed about the precise nature of the technology, its function and the content of the survey. Data collection was complete at the point at which consent to view and use the data was requested, and at this point participants were informed again that they may withdraw their consent at any point[4]. At the end of the study participants were given a debrief sheet (Appendix 2) containing details about the nature of the study and how their information is to be used, to give a level of transparency about the research being undertaken. This is presented at the end of the study as not to bias participants' method of interacting with the OLM during data collection. Contact details of the researcher were included should the participant have any subsequent queries.

Liaison with course lecturers prior to the implementation of the technology ensured that the academic content of the system was consistent with that of the School. This confirmed that the content remained appropriate, attainable and accurate for the participant students, and that the content should not cause any level of distress (e.g. undue difficulty in completing questions) or deception (e.g. through inaccuracy of the inferences generated). The aim of the technology is to show the learner an accurate assessment of their current state of

[4] No participants have indicated that they wished to withdraw from the studies during or after.

understanding of areas already covered in their academic programme, and so no undue risks are identified for students engaging with the content of their open learner model.

5.8.4 Conduct of the Analysis

Data captured for the purpose of analysis is anonymised in the logs, and throughout analysis is referred to only by sample number. No participants are uniquely identifiable in the analysis and the only link between the participant and the sample is information on the consent form. Analysis completed is consistent with the information given to participants at the time of evaluation and is used only to answer the research question posed in this thesis. As considered in Section 5.8.2 , no analysis is completed based on participants' personal, demographic or academic profile.

5.9 Summary

This chapter has further considered the research question and has detailed the design and rationale for the evaluation. It has revisited the requirements highlighted in the literature review, considered aspects of the research related to ethics, and has identified the need for an open learner model to be designed and implemented in order to undertake the evaluation. The design of the technology is now reported in Chapter 6.

Chapter 6

AN OPEN LEARNER MODEL IMPLEMENTING A DRILL DOWN APPROACH

The literature on metacognition, open learner modelling and information granularity/drill down approaches (Chapter 2 – Chapter 4) and the research method outlined in Chapter 5 require the design and implementation of a suitable piece of OLM technology, which should implement the drill down design detailed in Section 5.6.4 . This technology is specifically implemented as part of this thesis and is now overviewed in Section 6.1 .

6.1 System Concept

Literature suggests that providing students with only a black box representation of their learning is not sufficient (see Sections 2.3 and 3.3) and that technology such as OLMs help to address this. OLMs can provide educational benefits to the learner in terms of metacognitive development and improved self-assessment ability, and also support students in self-directed learning. OLMs permit students to decide on what elements of a course they wish to focus, when to acknowledge, receive and action feedback, and indeed on what level of abstraction they wish to work (e.g. everything on the course at once versus specific concepts and articles of information) (see Section 2.4). In addition there may also be need for students to access reference information, or information to justify particular inferences that that technology makes, in addition to observing the structure of information within a course, its domain content or specific detailed articles of information the technology may hold about a learner (see Sections 3.3 and 3.6).

In terms of OLMs, these key areas of design are also combined with the need for the OLM technology to support a wide variety of student strategies and cater for individual differences and different learning approaches. This thesis therefore proposes a design for an OLM that facilitates students' non-guided use of it in an educational setting. This is with particular reference to support that a drill down approach may provide in allowing learners to optionally access more fine grained elements of their learner models. The design is first expressed using the SMILI OLM framework standard (SMILI: Bull and Kay, 2016) (Section 6.2) and then is considered in terms of an underlying model (Section 6.3), methods of maintenance (Section 6.4), and its visual form (Section 6.5).

6.2 SMILI Framework Specification

Bull and Kay (2007, 2016) propose the Student Models that Invite the Learner In (SMILI) framework as a means to consistently describe open learner models, in the context of the reasons for opening the learner model to the learner (and other stakeholders) and different dimensions of the approaches taken (Table 10 and Table 11).

For investigating the use of a drill down approach as an effective means for the informational structure of the open learner model, this thesis considers Sections 1 (extent of model accessible), 2 (similarity to underlying presentation), 5 (access to sources of input), and 7 (presentation) to potentially be of the most relevance (covered in further detail below). The research question posed does not necessarily mandate the investigation of issues related to interactive maintenance, collaborative working or interaction paradigms beyond inspection, as a first port of call, but more of a formative assessment resource, highly suited to the self-directed learning practices.

Table 10: SMILI Framework description, Part I (x=important; xx=very important).

Elements	Properties	Accuracy	Reflection	Planning/ Monitoring	Collaboration/ Competition	Navigation	Right of access, control, trust	(Formative) Assessment
1. Extent of model accessible	Complete	xx	xx	xx		x	x	xx
	Partial							
	Knowledge level	x	xx	xx		x	x	xx
	Knowledge	xx	xx	x		x	xx	xx
	Difficulties	xx	xx	xx		x	x	xx
	Misconceptions	xx	xx	xx		x	x	xx
	Learning issues							
	Preferences							
	Other							
	Other users' LM							
2. Match underlying representation	Similar	xx	xx	x		x	x	xx
3. Access to uncertainty	Complete							
	Partial	x	x	x		x	x	x
4. Role of time	Previous							
	Current	xx	xx	xx		x	x	xx
	Future							
5. Access to sources of input	Complete							
	Partial	x	xx	xx		x	x	xx
	System	x	xx	xx		x	x	xx
	Self							
	Peer							
	Instructor							
	Other							
6. Access to model effect on personalisation	Complete							
	Partial							
7. Presentation	Textual (i.e…)	x	xx	xx		x	x	xx
	Graphical (i.e…)	x	xx	xx		x	x	xx
	Overview	x	xx	xx		x	x	xx
	Targeted/all details	x	xx	xx		x	x	xx
	All Details	x	xx	xx		x	x	xx
	Support to use							
8. Access method	Inspectable	xx	xx	xx		x	x	xx
	Editable							
	Addition	x	x	x		x	x	x
	Student persuade							
	System encourage							
	Negotiated							
9. Flexibility of access	Complete							
	Partial	x	xx	xx		x	xx	xx

Table 11: SMILI Framework description, Part II (x=important; xx=very important).

Elements	Properties	Accuracy	Reflection	Planning/ Monitoring	Collaboration/ Competition	Navigation	Right of access,	Assessment
10. Access initiative comes from	System User Peer Instructor Other	x	xx	xx		x	xx	xx
11. Control over accessibility (to others)	Complete Partial							
	System Peer Instructor Other							

With consideration of the reasons for which the learner model is opened, as given as headings in the framework, arguably those of learner reflection (as a key metacognitive activity) of assessment (which can considered here to be formative or of student self-assessment type) and of planning or navigation (such as best suits self-directed learning) are of the most relevance. The right of the student to access information held electronically about them is always important (and mandatory if they are to use it for metacognition) and the accuracy of the model is also relevant, but potentially not as great a priority as other elements. Indeed, the accuracy of the model is an initiation point for reflection, and studies (e.g. Bull and Britland, 2007) have shown that learners are able to detect inconsistencies.

It is a key part of this specification that the learner has the ability to ascertain (on demand) the information required for their current cognitive task, while the presentation of the model remains clear, simple and suited to purpose. This thesis considers students working independently in terms of self-directed learning, and so collaboration and competition is not considered as core to the research, although it can very naturally occur in groups of peers.

In terms of each of the principal elements of the framework in turn:

1. Extent of the model accessible. In considering a drill down approach, as much of the learner model as possible should be accessible, upon demand, however not all content should be immediately available. Following Shneiderman's (1996) approach, and those implemented in existing open learner models (e.g. Van Labeke et al., 2007) an overview is presented first with additional levels of detail available through navigation, therefore this may be considered *complete* access of the model, insofar as the learner model can be opened in a user interpretable form. The opening of knowledge level, knowledge and difficulty elements will be sufficient to provide the variance in granularity required to consider the approach from self-assessment, informational and utility aspects of the drill down approach in the research. Other learning issues and those pertaining to collaborative interaction are beyond the scope of the investigation.

2. Match to the Underlying Representation. The visualisation of the learner model must be interpretable by the user. For this reason, a direct representation of information persisted in database tables is not presented to the learner. The levels of aggregation of the information that are needed to produce information of the coarser grained aspects of the visualisation of model, or partial models, mandate that this is different.

3. Access to Uncertainty. Uncertainty is not required to be explicitly presented to investigate the use of a drill down approach. However, user interpretation of the model at coarser grain sizes may lead learners to identify elements of uncertainty within the data, such as incomplete concepts within the model, or conflicting information at belief level. For discussion of the visualisation of uncertainty, see Demmans Epp and Bull (2015).

4. Role of time. The primary aim of an open learner model is to make visual a prediction about the current learning state and needs of the learner. In investigating a drill down approach this thesis considers information that is a current depiction of the learner for their purpose of planning, reflection and self-direction in learning. Arguably, the presentation of information such as previous learner model states (such as Ginon et al., 2016) may also provide additional fine grained information, however this is not key to the evaluation of the research question.

5. Access to Sources of Input. All information in the learner model originates from inferences made about the learner that result from their interaction with the system. This presents the system as an inspectable resource. Adding a further layer whereby the student may be able to see specific questions asked and the answers given would allow a complete level of access to sources of input. The investigation of this thesis, however, focuses specifically on the content of learner model inferences in the initial case.

6. Access to model effect on personalisation. This aspect of an open learner model is not core to an investigation that relates to information structure, however it may become so should research focus on aspects of learner model accuracy.

7. Presentation. Presentation is of particular relevance to this research. At all levels an open learner model should strive for simplicity (see also the examples of Section 3.5). The approach of this thesis considers starting with an overview, which may be broken down into finer elements of the model, including informational structure, and then specific beliefs that are inferred. Drawing the user's attention towards specific elements at each part of the process, relevant to cognitive task, leads to a definition of the use of targeted detail (e.g. this topic has x concepts). The approach considers making all details of the model available to the learner (e.g. weights in the model), in a structured way, making use of graphical and textual information.

With regard to support to use, there is no active component intended for this that might bias interaction or disrupt user flow or self-direction. However, necessary elements are required for the interpretation of the visual methods, such as keys, where colours are used, and potentially also textual descriptions, where there are domain specific representations that might otherwise require prerequisite skills to be able to interpret. This thesis also considers the opening of the domain model alongside belief descriptions to be an element of support to use, as they provide a reference point against which the learner may interpret their own belief.

8. Access Method. The approach required considers that of an inspectable learner model. While other approaches of interactive maintenance may be combined with it (e.g. Van Labeke et al., 2007), an inspectable-only approach is sufficient for the research. The learner should have the option to add evidence to the model via a suitable means (e.g. multiple choice questions) to allow inferences to be updated.

9. Flexibility of Access. In accessing the open learner model in a drill down manner there are options regarding how to access the model content. These mainly include the level of granularity of some of the information, and the areas of the domain that are inspected in greater detail.

10. Access initiative comes from. In all cases the access initiative comes from the learner. It is they who initiate access to the model and its various components, as best suits their short term informational requirement or cognitive goal.

11. Control over accessibility to others. There are no other stakeholders considered; the evaluation of the research question concerns the learner working independently in a self-directed learning context.

Following on from the consideration of the architecture in terms of the SMILI framework, this thesis now considers the issue of learner modelling in more detail.

6.3 A Simple Learner Model

The evaluation of a learner modelling procedure is not core to the aims of the research. It should therefore only be as sophisticated as is required to support the modelling and display of information on different levels of granularity, such that a drill down approach can be facilitated. It must, however, be as fine grained as the finest component the open learner model need display. It must also provide suitable formative assessment of the learner.

Figure 4: model pipeline, per concept.

Following simple weighted modelling approaches (as taken by Mabbott and Bull, 2004; Johnson and Bull, 2009; Mabbott and Bull, 2006), and balancing the need to disregard or compact older information (e.g. Kay, 1994), it is proposed to implement a simple pipeline based approach, taking the last 4 inferences of student understanding for each concept, and weighting these so that the most recent receives the most influence, and so on (Figure 4). The content of specific beliefs should come from the ability to apply (or mis-apply) beliefs and rules from the domain model (see Section 3.2).

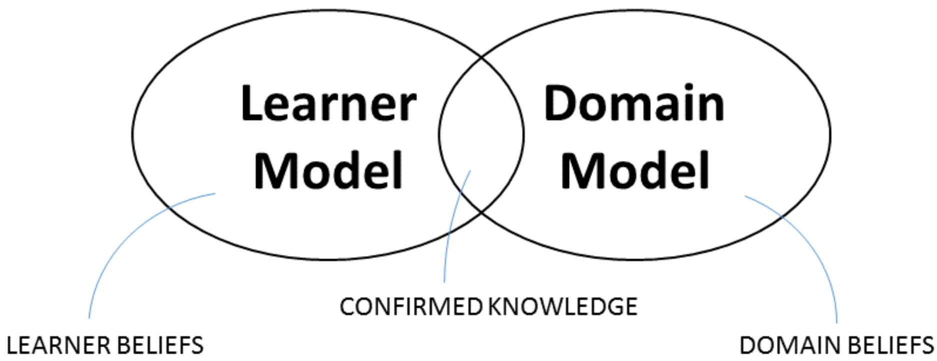


Figure 5: perturbation model.

Where the learner and domain beliefs match, they may be combined. If a learner belief matches a domain belief for the concept, this is confirmed knowledge; where learner beliefs do not form part of the domain model, they may be considered as problematic knowledge. It is important to note that there may be multiple correct beliefs for a concept. This implements a perturbation model approach (Figure 5).

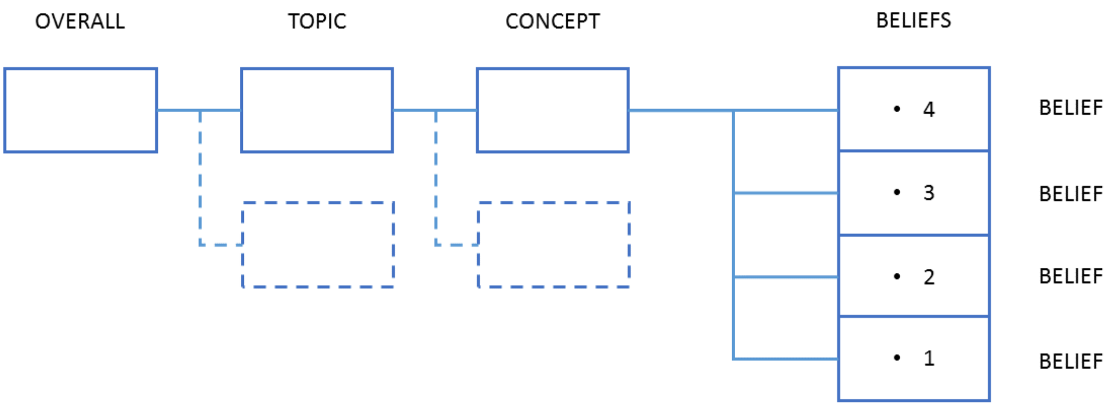


Figure 6: aggregating information.

Thus, opening the learner model on the finest level of granularity should show the specific beliefs (applied and demonstrated knowledge) from the last four pieces of information (far right of Figure 6). Moving leftward in the figure, each belief has a knowledge category associated with it (confirmed or problematic) to state whether it matches or contradicts the domain model. The next coarsest level of information is to display a breakdown of the combined knowledge levels, while also showing the weightings of the information. (e.g. if the newest two beliefs were confirmed knowledge, the third problematic, and there was no fourth piece of evidence, the learner model would state 40%+30%=70% probability it is understood; 20% not understood; 10% do not know). The next coarsest level could then aggregate this into a single value, and the next level would be an average of all concepts in the topic to give the model for the topic (partial model), and again the overall value is an average of topics. This mirrors the structure of information as outlined in the state model of Section 5.6.4 .

6.4 Updating the Learner Model

The modelling procedure requires inferences about student understanding and the model can be updated at any point. In this section this thesis looks at suitable methods for capturing information about student understanding and suitable domain content.

6.4.1 Domain Content and Methods for Obtaining New Information

To provide beliefs that can be used by the modelling process, information needs to be collected from the learner about their understanding. This is not part of the model, nor part of the modelling process, but in terms of providing continuity of action for the learner it is intended to make it available within the same piece of software. Among a variety of suitable data sources (such as those from quiz results, class exercises or multiple choice questions, exams; see Section 3.2), the use of multiple choice questions is selected as a simple method, suited to purpose. In the STEM domain, and in question scenarios where the student is asked to apply a rule or demonstrate the application of a technique, selection of an answer from the options list can be used to ascertain correct understanding, if the correct answer is selected.

Distractor answers can be generated by applying common misapplications of the rules or techniques. It is the rules or techniques applied, as indicated by the student's selected response, that may be passed as evidence to the learner model. (See Section 6.4.2 for question and domain development.)

Furthermore, to assess issues related to student self-assessment accuracy, measures of this need to be taken at suitable points during interaction (thus the system models the students' behavioural and diagnostic information together, to the same precision). It is suggested that to make this directly comparable to the state of the model, this should be built into the mechanism for responding to questions; this is can be modelled with the same precision as for knowledge. Following approaches taken by Bull and Pain (1995) and Kerly et al. (2008) this is done using a 4-point Likert scale at the point of submitting the answer ("not confident" to "very confident"). Students are required to state their confidence in terms of being able to respond correctly to the question, as a self-assessment (diagnostic model). The system's assessment is that of whether the student was able to correctly respond to the question (behavioural model).

In the STEM domain and at university level a core foundation of many engineering based courses involves a common set of mathematical principles and elements of discrete logic (such as Boolean). Open learner models have also been implemented in this, and similar domains, (see Section 3.2), and technology enhanced methods are appropriate for mirroring tutors' marking practices in the mathematics domain (Ashton et al., 2006, in Ras et al., 2016). Within the department of study (Electronic, Electrical and Systems Engineering, University of Birmingham) there are a good number of students who have (and will) cover basic engineering mathematics and Boolean logic aspects as part of their degree (see Section 5.4). Open learner models have previously been implemented at university level (e.g. Mitrovic and Martin, 2007) and OLMlets (Mabbott and Bull., 2004) has been previously used in the school with domain content including basic engineering mathematics and Boolean logic.

The system in this thesis builds on an established question corpus, and course design, to provide the domain content for evaluation. In both domains (basic engineering mathematics; Boolean logic) multiple choice questions had already been designed by the course instructor and were used with other technology and open learner model systems. Each question asked for a mathematical or logic problem to be solved and provided several specific answers. The information passed to the learner model was just knowledge level information (i.e. indication that the student can apply domain knowledge with regard to a given concept, rather than what the domain knowledge was that was applied). In order to investigate the drill down approach, finer grained information is needed, to state what the articles of knowledge are that are applied. A further criticism of this approach by course lecturers has been its static nature, that a small number of the questions allowed students to memorise answers and that students were observed attempting to game the system (i.e. as described in Baker et al., 2013). One further consideration is that while open learner models, and indeed the approach described here, can be to some extent domain independent, it is necessary to note that the technology used to generate the questions and elicit learner beliefs must have application of domain content. This is the case whether the application of the domain content is part of a domain model or embedded to a certain extent in the design of the questions, or a question generation mechanism. This is now considered in the next section.

6.4.2 Question Design and Question Templates

Building on the above requirements of the learner model (Section 6.3) and the method of updating the learner model in the context of domain content (Section 6.4.1) the approach of this thesis considers the use of automatic question generation from templates. The templates have knowledge of domain content embedded within them and also contain common examples of the misapplication of mathematical/logic rules to provide distractor answers. Both aspects of this information are available to provide the modelling system with enough information to be able to open the learner model and corresponding elements of the domain model to the learner, as the finest level of detail in the drill down capabilities of the interface.

Luecht (2013) suggests that this may be done in a top down approach (starting with constructs, tasks and models of the domain to be covered) or from a bottom up approach (starting with existing resources, established and tested in the domain). This thesis takes the latter as a starting point and establishes 128 templates for the basic engineering mathematics domain, which is identified as comprising 5 topics and 26 concepts, and likewise the Boolean logic course content is made up of 3 topics and 9 concepts, mirroring the delivery students would receive in the course. Using the template approach contributes strongly to an aspect of automation, supporting teachers in intensive tasks, and is highly reusable.

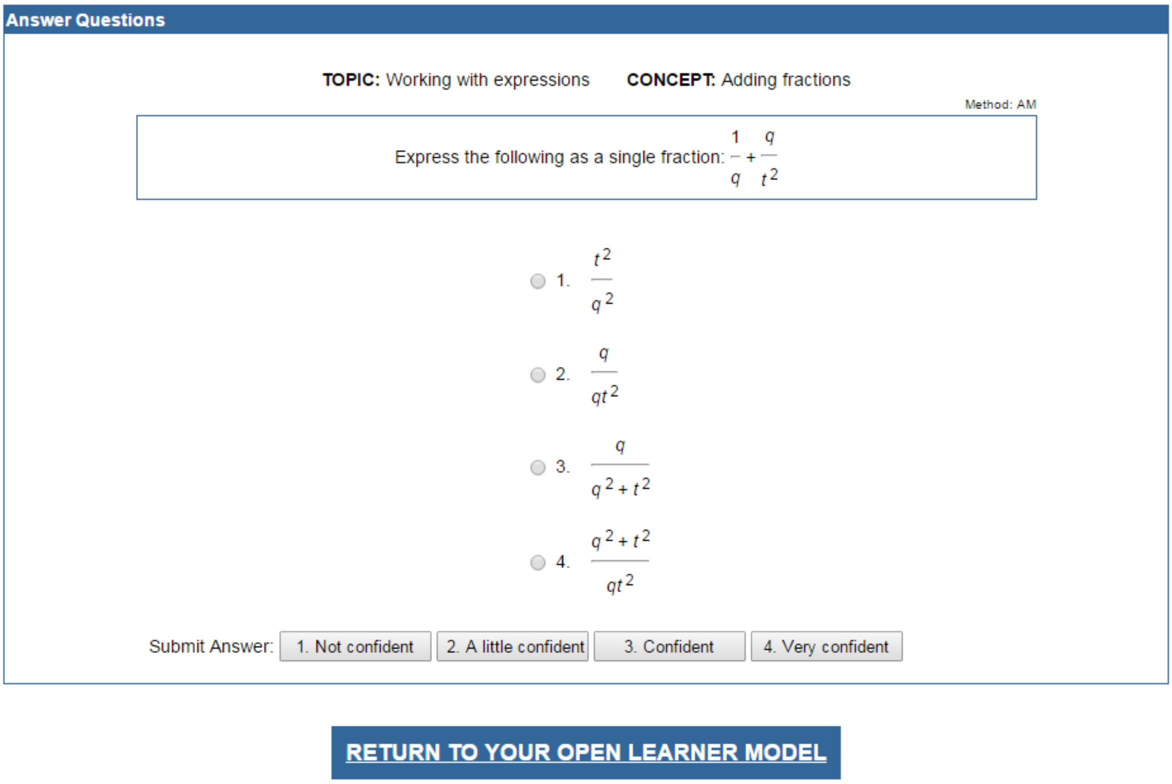


Figure 7: question interface.

For each template correct mathematical methods are applied, as would be part of the domain model, to give the correct answer for a given problem (e.g. addition of fractions, as shown in Figure 7. Correct answer – Option 4). The numbers required by the variables are randomly generated, within tolerable ranges, to give achievable answers that students may calculate at the difficulty level required by the course. Most templates require the application of only one mathematical method, although several are compound.

A library of mis-applied and misconceived methods was developed, in collaboration with the course lecturers, to provide distractor answers, and beliefs to be inferred that are outside the domain model (see perturbation model – Figure 5). In the example shown in Figure 7, the correct application of the rule is to sum the multiple of the leading diagonal (1 x t^2) and back diagonal (q x q) to gain the numerator, and to multiply the two denominators (q x t^2) to get the denominator – giving Option 4 ($(q^2+t^2)/(qt^2)$). An example misapplication of the rule may be to multiply the leading diagonal to get the numerator ($1 \times t^2$) and the back diagonal to get the denominator ($q \times q$), thus option 1 (t^2/q^2). The rules used to generate the specific question response given are the articles of information that are stored in the learner model.

The question interface (Figure 7) states to the learner the area of the domain (topic, concept), the question, a range of answers, and the options to either submit the question or return to the learner model. The next question is automatically presented, thus it is the learner's decision to break away from this to inspect the learner model, and likewise, when the learner returns to the learner model, the same question is presented, thus the learner may also use the OLM as a point of reference before commencing a task if they wish (in alignment with Whitelock's (2011) Advice for Action approach, in addition to following best practice for facilitating formative assessment opportunities (see Chapter 2 for further commentary).

6.4.3 Learners' Selection of Question Content

To mirror each of the ways by which students were able to engage with the course content, and with existing resources, students are free to decide whether to answer questions on very *specific concepts* within the course, any concepts related to a *specific topic* within the course, or *all topics/concepts* at once. As per the self-directed learning approach (Section 2.4), the student, rather than the system, remains in complete control over when information is accessed and when new information for the model is provided. The system will as such indefinitely allow the student to answer questions on their chosen focus, until a decision is taken by the student to change focus or inspect learner model information.

6.5 Opening the Learner Model

Opening the learner model to the learner is of key importance in the evaluation. Several issues are to be considered:

- The arrangement of different information types on the interface, with view to facilitating a drill down approach through mouse interaction. (Section 6.5.1)
- Appropriate visual methods for opening the learner model at each level of granularity, supporting relevant cognitive tasks. (Section 6.5.2)
- Provision of sufficient levels of contextual information to allow the interpretation of components of the learner model that are externalised. (Section 6.5.3)

6.5.1 Arrangement of the Interface

There are 4 main levels of granularity to be considered here as part of a drill down approach (see Table 8, Section 6.1). From the perspective of analysis, at each level the selection of an area in which further information will be presented should be traceable, for example in terms of a mouse click or navigational event, such as indicates the move from inspection of one element to the inspection of the next (as outlined in the system overview, Section 6.1).

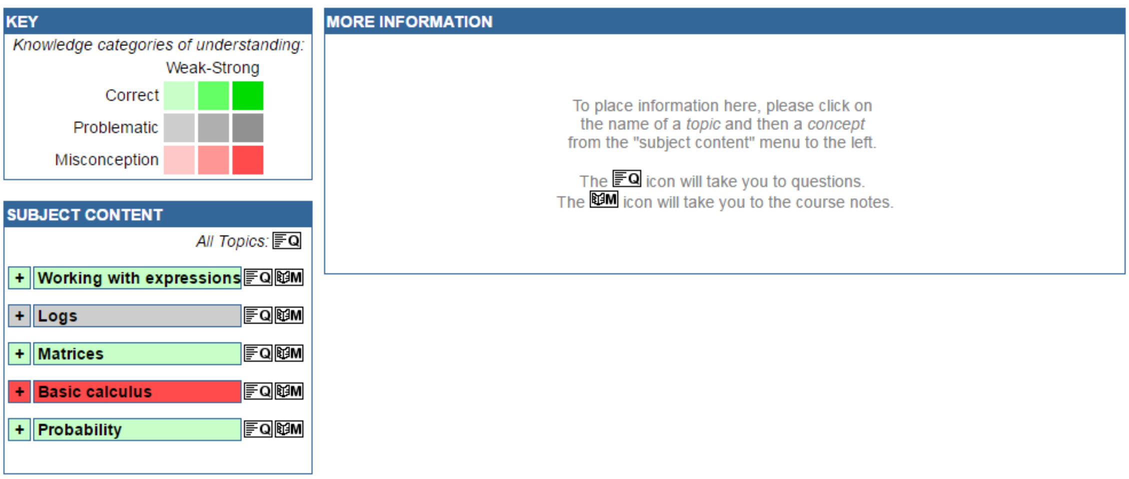
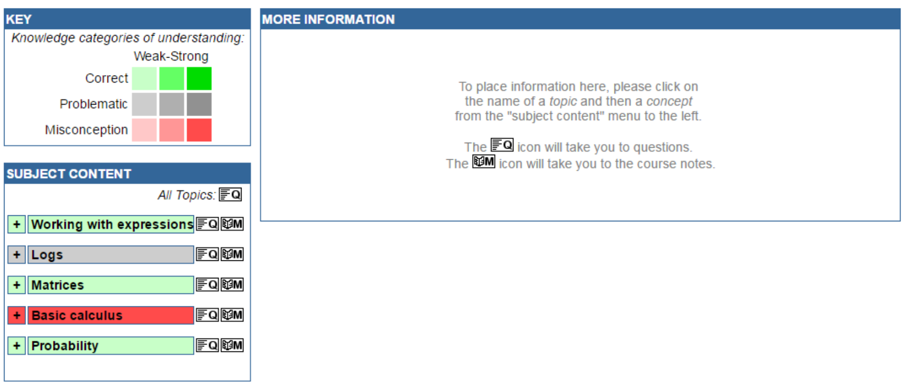
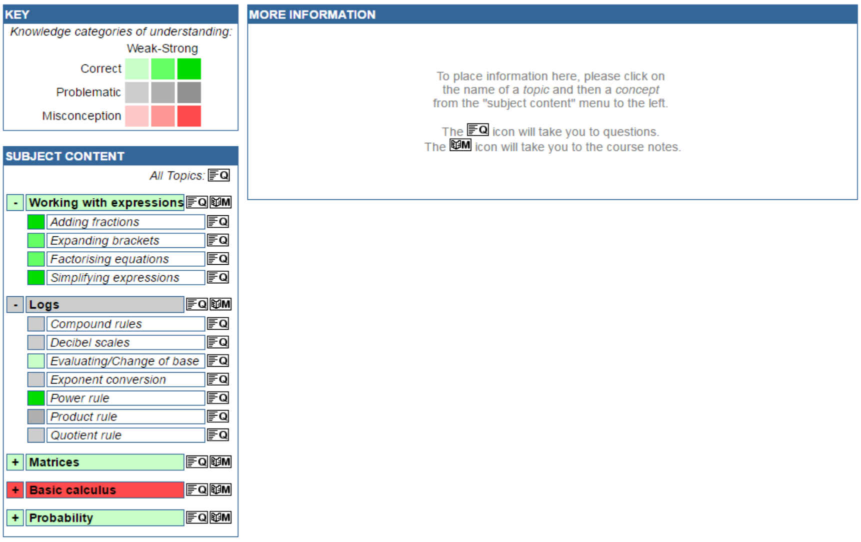
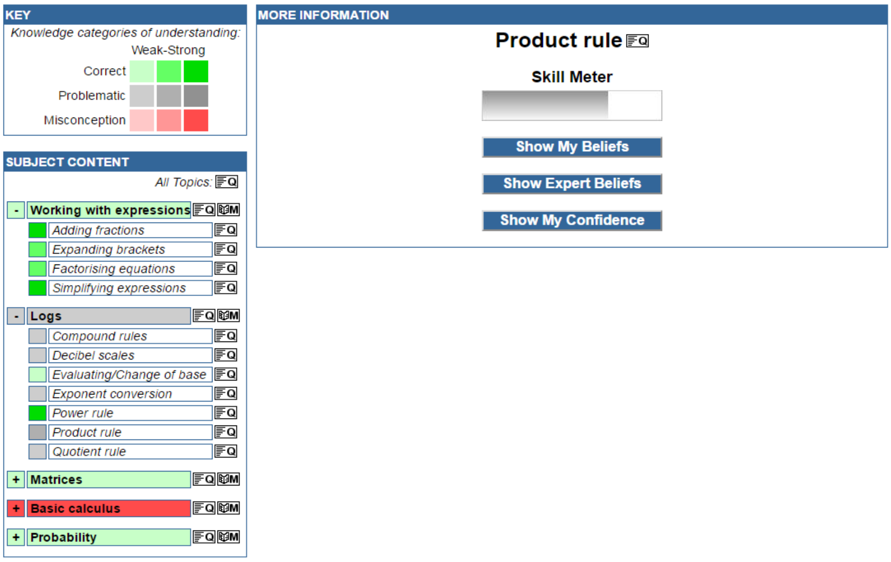
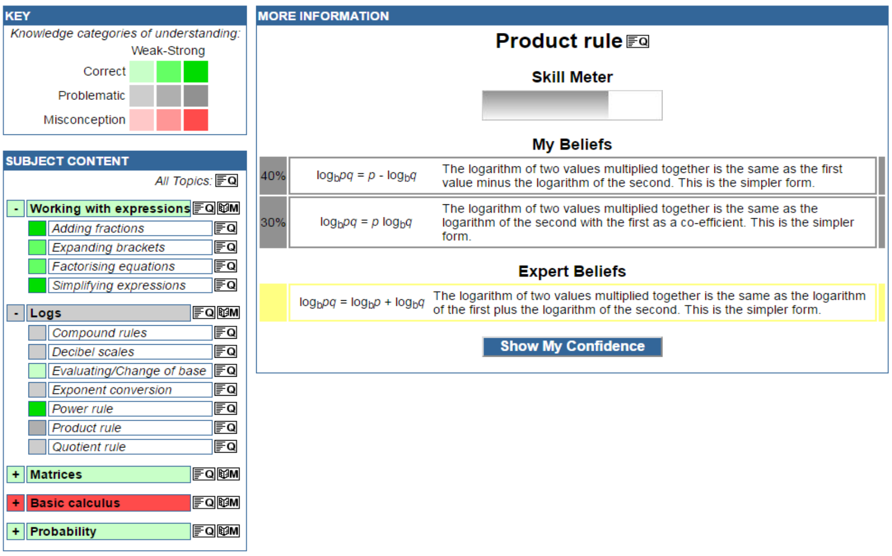


Figure 8: OLM interface - high level.

Thus, this thesis considers an arrangement whereby the coarsest level of information is presented initially (topic overview – left of Figure 8) and through mouse interaction the breakdown of sub-components (concepts) can be revealed (e.g. as in Kay, 1994). The next levels of the drill down approach, which describe the breakdown of information within the concepts, initially by knowledge level (showing part of the modelling) and specific beliefs, will take up a greater amount of screen real estate, and so have an area of the interface designated for this (right of Figure 8) (e.g. as in Van Labeke et al., 2007, Johnson et al., 2013b). For a specific concept which is selected for drill down, the information at all levels of granularity will be able to be displayed concurrently, and the high level overview will remain present to allow other navigational possibilities. Elements of the drill down approach are kept in the same area of the interface, for consistency, and so that information of a specific type can be associated with a visual area. For this reason this system does not use forward or backward navigation, or modal dialogues. Drilling down into the learner model will thus constitute an interaction sequence as shown in Table 12.

Table 12: example drill down in to the open learner model. Cross-reference Figure 2.

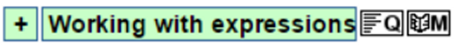
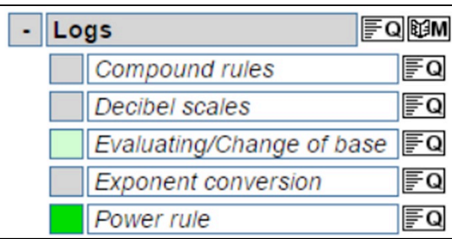
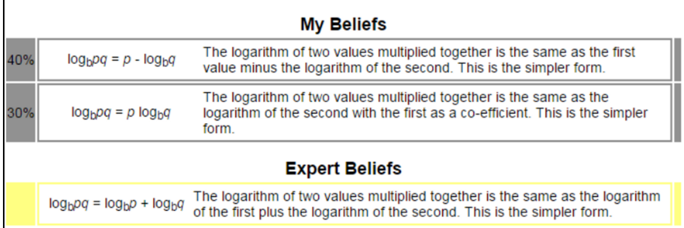
Level	Example
Overall. Summary of all topics is shown on the left of the screen	

<table>
<tr><th>Level</th><th>Example</th></tr>
<tr><td>Topic level breakdown. Clicking on a specific topic reveals the concepts of which it is composed.</td><td>
</td></tr>
<tr><td>Concept level breakdown. Clicking on a specific topic shows information about the concept in the area on the right. Shown is a breakdown of the knowledge level in terms of the model weightings (see Figure 6, Section 6.3).</td><td>
</td></tr>
<tr><td>Belief level breakdown; domain level information. Clicking on the relevant labels opens the specific beliefs held about the learner, together with a reiteration of their contribution to the model in terms of weights.</td><td>
</td></tr>
</table>

6.5.2 Visual Methods for Opening the Learner Model

At each level of abstraction the learner model should be opened in a suitable form, using established techniques. The open learner model should strive for simplicity at all levels (Uther and Kay, 2003). While using only one visual method at each point of entry (as opposed to multiple isomorphic learner model visualisations – e.g. Mabbott and Bull, 2004; Johnson et al., 2013b) may be restricting in terms of supporting user preferences and individual differences, the work of this thesis is more concerned with the use of the drill down technique, and therefore aims to reduce the number of variable factors in terms of learner interaction and possible navigational paths (restricting them to those detailed in Figure 2, Section 5.6.1). Building on the need for simplicity, the following are proposed, as in Table 13, as appropriate visual methods.

Table 13: visual methods for opening the learner model.

Level	Method	Example
Overall	Coloured nodes (e.g. Kay, 1994)	
Topic level breakdown	Coloured nodes, in a tree structure (e.g. Mabbott, 2009). Hierarchical relationship shown.	
Concept level breakdown	Skill meter, to show how much of each knowledge category is present, according to model weightings (e.g. Mitrovic and Martin, 2007)	
Belief level breakdown; domain level breakdown.	A domain specific description of the learner beliefs, i.e. shown in mathematics notation. Textual descriptions are good practice to explain the content of the mathematics description. Domain model information is available for contextualisation (See Section 3.6)	

6.5.3 Provision of Contextual Information

While it is not proposed that an active amount of support to use is implemented (e.g. overlays, interactive tutorials), this can be facilitated though a sufficient introduction of the system to the participant at the start of the student's evaluation, and the availability of teaching assistant support during laboratory sessions. In terms of the learner's interaction with the system it is however important that the necessary information is available to interpret the open components of the learner model, and that navigational possibilities are clear. A key is included for the colour coding of elements representing knowledge level (see top left of Figure 8), and where mathematical descriptions are present (e.g. as in Table 12), corresponding explanatory text is also included. Likewise, where specific beliefs are shown, the domain's understanding is also available for inspection to provide the additional context in which to interpret beliefs and compare and contrast understanding. (At coarser levels of granularity, i.e. knowledge level, this is implicit, as the domain should be expected to be 100% correct).

6.6 Summary

This section has presented design considerations and the implementation of technology suitable to investigating aspects related to use of a drill down approach in an open learner model and implements the interaction state model specified in Section 5.6.4 . The thesis now moves forward to report on the evaluation of this technology with end users (Chapter 7 and Chapter 8).

Chapter 7

STUDY 1

An evaluation of the research question with participants is split across two studies. This chapter reports on Study 1, according to the plan outlined in Chapter 5, and presents analysis and commentary on the sub-research questions as outlined in Section 5.1 . As suggested by the literature, it first considers user acceptance of the technology of those who volunteered to participate in the study. This chapter then proceeds to provide a description of how drill down is used within the OLM before reporting on its impact on self-assessment accuracy and tasks relating to the regulation of cognition. The experimental groups are as described in Section 5.6.1 .

7.1 How is an Open Learner Model with a Drill Down Approach Used?

This sub-question is considered from the perspective of consistency of use, user acceptance of the participants who decided to engage and the depth to which the model is used. This gives a description of students' behavioural interactions with the OLM. The aspects analysed as part of this sub-question are used to contextualise the impact on self-assessment accuracy and support for tasks regulatory to cognition, which are considered later in the chapter.

7.1.1 Is an open learner model with a drill down approach accepted by its users?

There was a good level of usage of the system (Table 14, Section (i)) with learners answering a mean of 63.75 (maths) or 37.92 (Boolean) questions where the model was uninstantiated. For the two groups where the model was instantiated on initial access, students continued to answer an average of 27.63 (maths) or 22.0 (Boolean) questions. The OLM therefore had a good amount of data to display. It was accessed regularly and on multiple occasions (average of MA(U) 22.5; BO(U) 9.23; MA(I) 11.75; BO(i) 9.2). Across conditions there is a slightly less intense interaction for Boolean logic than for mathematics. Where the model had an initial import of data, participants continued to inspect and update the model for the remaining time in the lab sessions. All participants made some use of the drill down functionality, although this was greater for some than for others, as is reflected in the larger standard deviations. The high and overlapping standard deviations would also suggest little statistical significance in terms of the consistency of student interaction, but this does establish that drill down is used.

Across all 4 groups the mean number of questions answered between model inspections is between 1.31 and 4.02, with medians in the range 1.08 to 1.35, which suggests a strong bias towards working with very short bursts of questions and regular OLM inspection for the majority of interaction.

Table 14: general usage statistics.

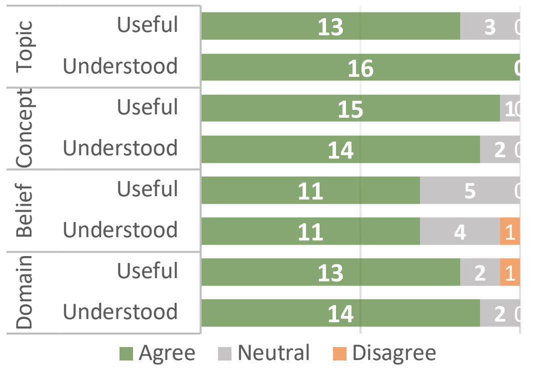
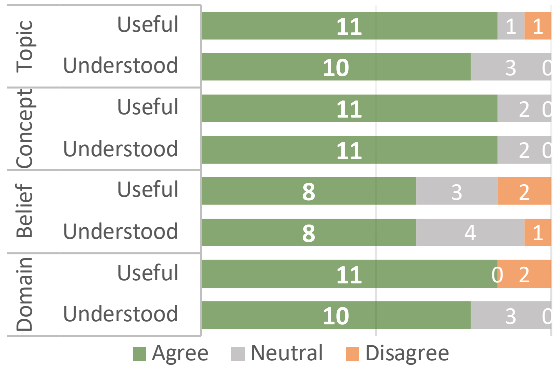
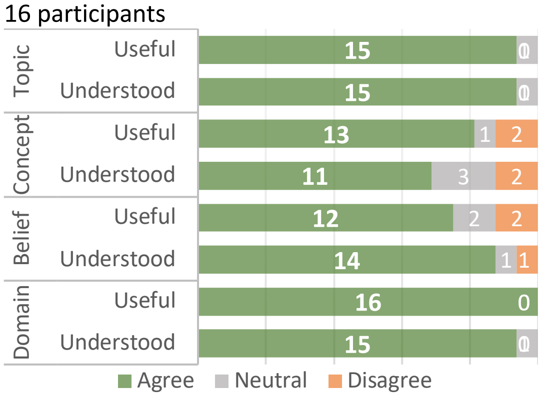
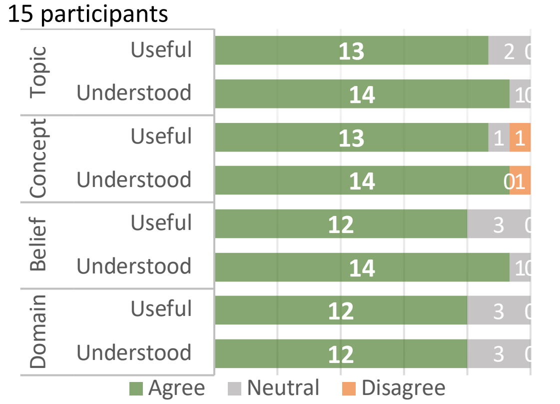
Usage		MA (U)	BO (U)	MA (I) [5]	BO (I)
(i) Number of participants		16	13	16	15
(i) Number questions answered		63.75 (SD 42.83)	37.92 (SD 22.81)	27.63 (SD 24.45)	22.0 (SD 21.95)
(i) Number of OLM accesses		22.56 (SD 12.48)	9.23 (SD 9.26)	11.75 (SD 12.52)	9.2 (SD 9.62)
(i) Average number of questions before OLM view	Mean	2.58 (SD 2.56)	4.02 (SD 6.17)	1.31 (SD 2.83)	2.45 (SD 6.62)
	Median	1.15 (SD 0.98)	1.35 (SD 1.43)	1.08 (SD 1.62)	1.11 (SD 1.05)
(ii) Number of inspections of more detailed model elements.	Topic	25.19 (SD 18.47)	10.46 (SD 6.11)	25.69 (SD 24.90)	14.87 (SD 12.59)
	Concept	25.19 (SD 18.19)	13.62 (SD 10.07)	31.38 (SD 33.96)	28.00 (SD 22.34)
	Belief	7.00 (SD 4.12)	3.92 (SD 3.15)	6.75 (SD 5.07)	6.40 (SD 4.34)
	Domain	5.5 (SD 3.46)	4.46 (SD 3.43)	6.38 (SD 4.43)	5.67 (SD 3.54)
(iii) Frequency of use, per model inspection	Topic	1.19 (SD 0.58)	1.49 (SD 0.69)	3.55 (SD 2.50)	2.68 (SD 1.77)
	Concept	1.26 (SD 0.90)	1.82 (SD 1.51)	5.10 (SD 4.38)	7.53 (SD 7.99)
	Belief	0.39 (SD 0.29)	0.78 (SD 0.62)	2.04 (SD 2.85)	1.83 (SD 2.10)
	Domain	0.32 (SD 0.30)	0.78 (SD 0.63)	1.99 (SD 2.83)	1.77 (SD 1.98)

[5] For datasets where participants had to answer questions before having access to the open learner model, these numbers reflect the level of interaction (i.e. questions answered, OLM accesses) after access to the OLM was gained.

Per model inspection (Table 14 – Section (ii)), the coarser grained elements (such as topic overview, concept overview) were accessed more frequently than the finer grained elements (such as specific beliefs, domain content), and the distribution is similar between the maths and Boolean domains. This is in part a reflection of the structure of the information, as it is required to access information on a concept before accessing a belief. This does however indicate that <u>not all drill downs were to the deepest level of inspection</u>. In considering the effect of the model being instantiated at the point of first access, the frequency of inspection is much greater in both domains, with the frequency of inspecting concepts being the greatest. This is potentially so as this is the level of granularity upon which the questions were answered. It is important to note at this point that each of the domains contained 3 topics, however there were 9 concepts to explore (and consequently 9 sets of beliefs and 9 domain representations).

In consideration of the frequency of inspection per question answered as a further measure of level of use (Table 14 – section (iii)) again the distribution is similar between the two domains, in both instantiated and uninstantiated cases. The use of the access of beliefs is of the lowest frequency. The frequencies are much higher for the two instantiated groups, potentially as there is much more information to explore upon any one navigation to the OLM.

Table 15: survey responses - whether intervention was understood/useful.

	Mathematics	Boolean
Uninstantiated	16 participants  	13 participants 
Instantiated	16 participants  	15 participants 

As part of user acceptance, on the post-usage survey (Table 15, Appendix 3) <u>the majority of participants indicated that they understood the representations of the learner model</u> at each level of granularity, in each group, with several disagreements in the MA(I) group. When considering the questions in pairs, the responses for "understood" generally correlate with "useful" in each case. Of the two mathematics groups, the belief level representation was the least understood/useful. However, when compared to the open representation of the domain model (same granularity level), the domain representation was understood to a greater extent. <u>There is potential here that students were struggling to interpret detailed information about themselves</u>. Of the two Boolean logic groups, there are also fewer "agree" responses than for mathematics, which might indicate differences in the complexity of the information in the domain. Likewise, there are a greater number of agree responses when the model was already instantiated, potentially because of the increased richness of information available about the learner from the initial interaction point.

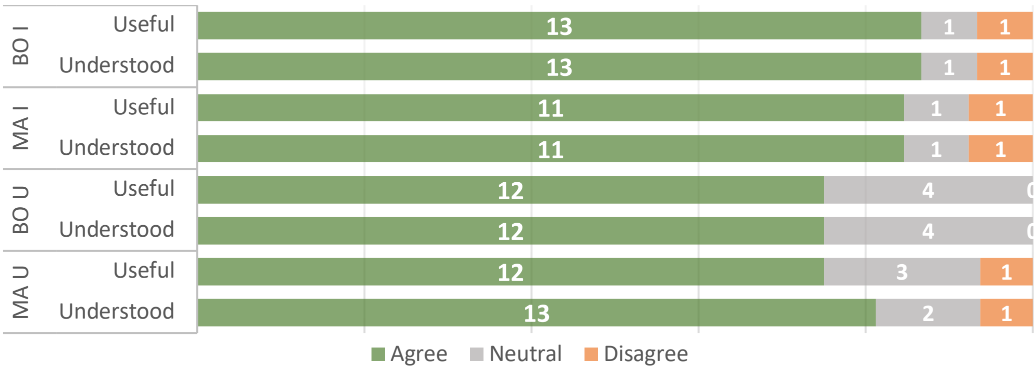


Figure 9: survey responses – whether drill down approach was understood/useful.

When asked about specifically about the use of a drill down arrangement of the OLM (Figure 9) students generally understood the drill down approach, in each of the 4 groups (6 of the total 60 participants disagreed) and that they found the approach useful. (This is slightly more the case with the two groups with instantiated models).

7.1.2 Is use of the drill down approach in the open learner model consistent across time?

In the normalised cumulative distribution frequencies (CDFs) presented in Figure 10 to Figure 13 the blue (answering questions) and green (access the OLM – high level) lines define the regularity with which the learner model is accessed. They show a fairly linear distribution in each case, indicating regularity of OLM access across time. These are the points of reference around which use of the different drill down components is made. In all cases, it is worth noting that the red (inspect beliefs) and purple (inspect domain) lines follow a very close pattern. This mirrors the description given in the state model (Section 5.6.4) whereby the inspection of a belief is more likely to be followed by the inspection of the corresponding element of the domain, above other navigational possibilities. Likewise, the orange (concept) and yellow (topic) lines also follow a similar distribution - these are the coarser levels of the drill down.

With reference to the two groups who used the model once it was initially filled with information, the majority of deep level inspections (purple, red, and then orange) occurred in the initial period following this. The linear slope of the other lines (yellow, green blue) indicate that interaction did indeed continue beyond the point of initial access and that learners proceeded to answer further questions to update the model and also continue to use the drill down approach beyond this initial period (although it is intensive in the initial period). Overall the distributions for coarser and finer granularity components do not follow exactly the same path, but with the display of a slightly positive exponent for finer granularity in the case of three of the groups. The information in the Boolean dataset is also noisier than the mathematics dataset, potentially because of the lower number of interactions over the same period. In all four groups, the purple (domain inspection) and red (belief inspection) lines follow a very similar distribution. This might further imply that these two resources are being used together. The yellow and orange lines (topic inspection, concept inspection) are proximate, implying that they may be being used for similar purposes. In the two experimental groups where the model was complete at the start of interaction slightly more use is made of these coarser grained elements earlier in interaction, as would be consistent with systematic overviewing of the model's contents. Where the model was initially empty, more regular use of these coarser grained elements takes place later in interaction when there are greater volumes of information to overview.

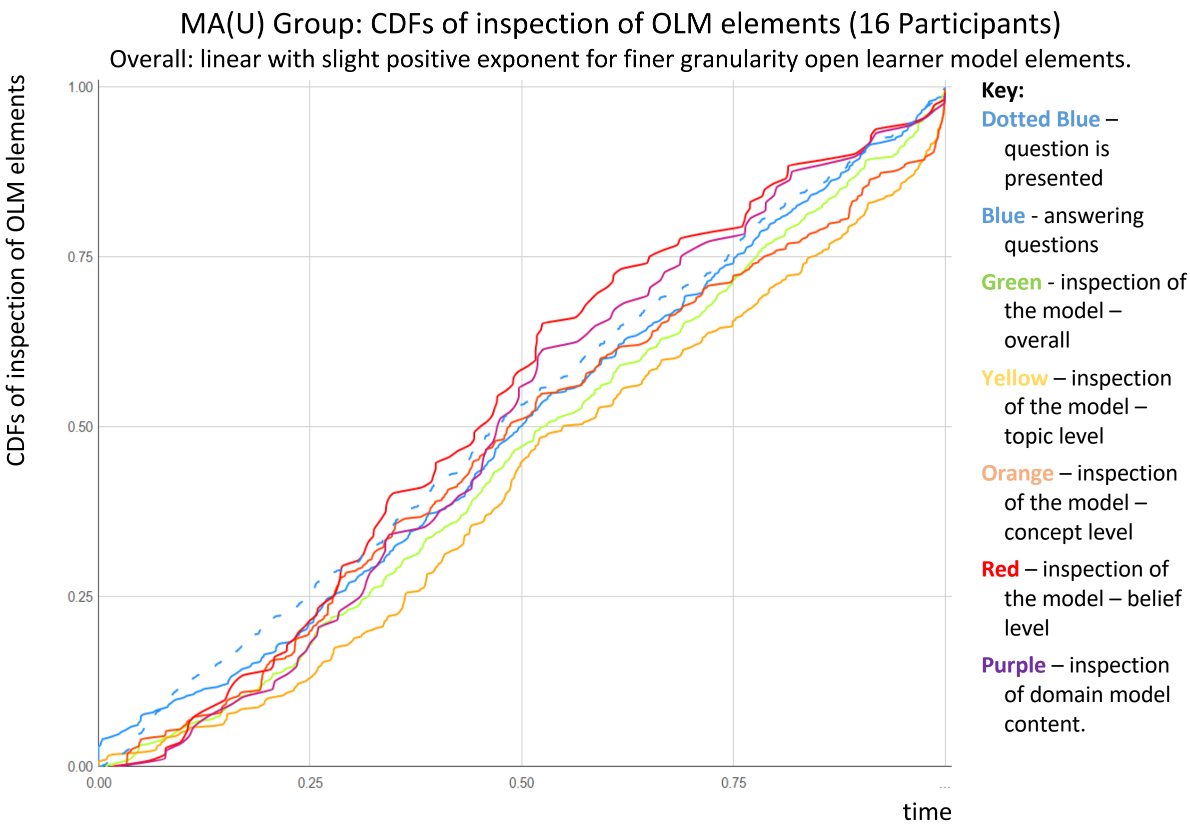


Figure 10: MA(U) group cumulative distribution frequencies of interface usage across time.

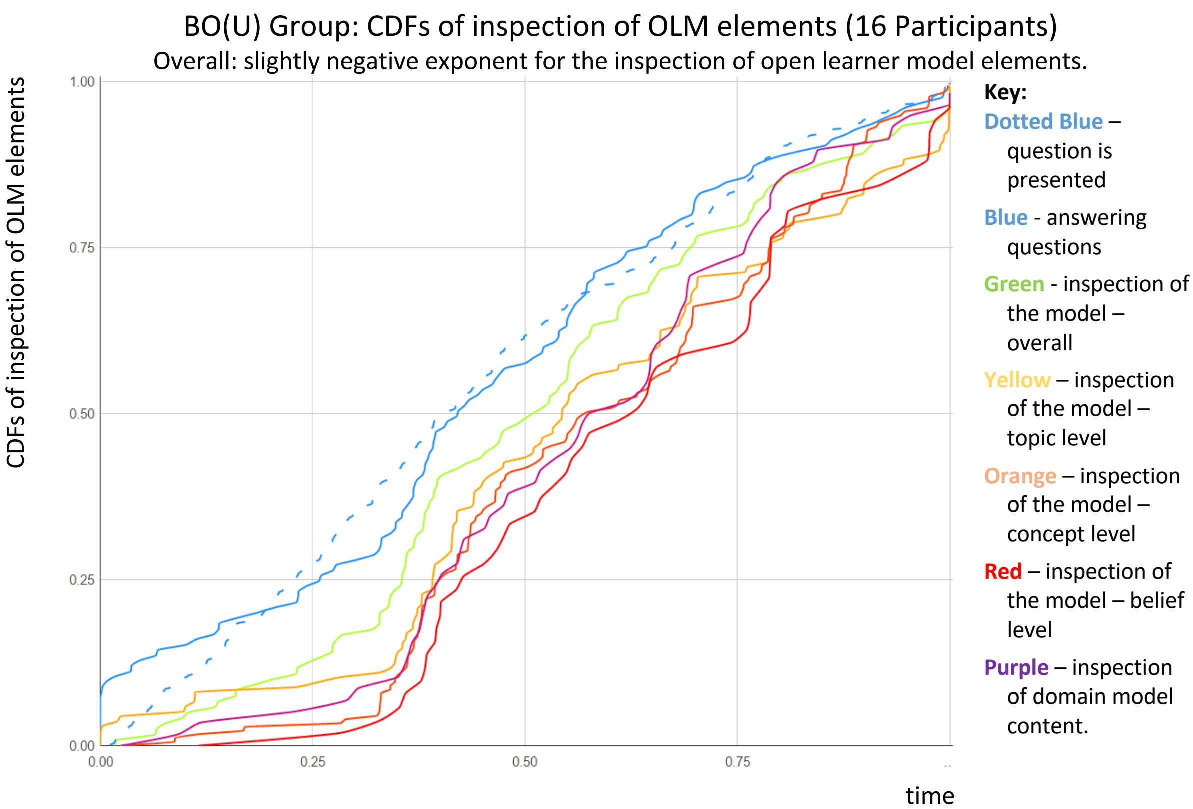
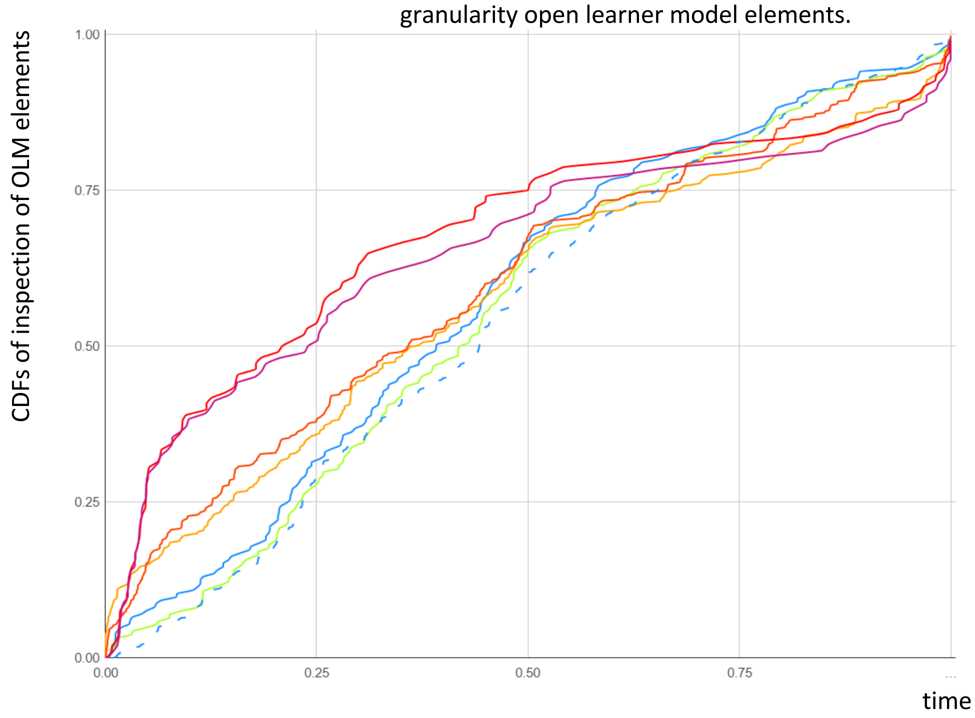


Figure 11: BO(U) group cumulative distribution frequencies of interface usage across time.

MA(I) Group: CDFs of inspection of OLM elements (13 Participants)

Overall: slightly positive exponent for interaction. Defined positive exponent for the inspection of finer granularity open learner model elements.

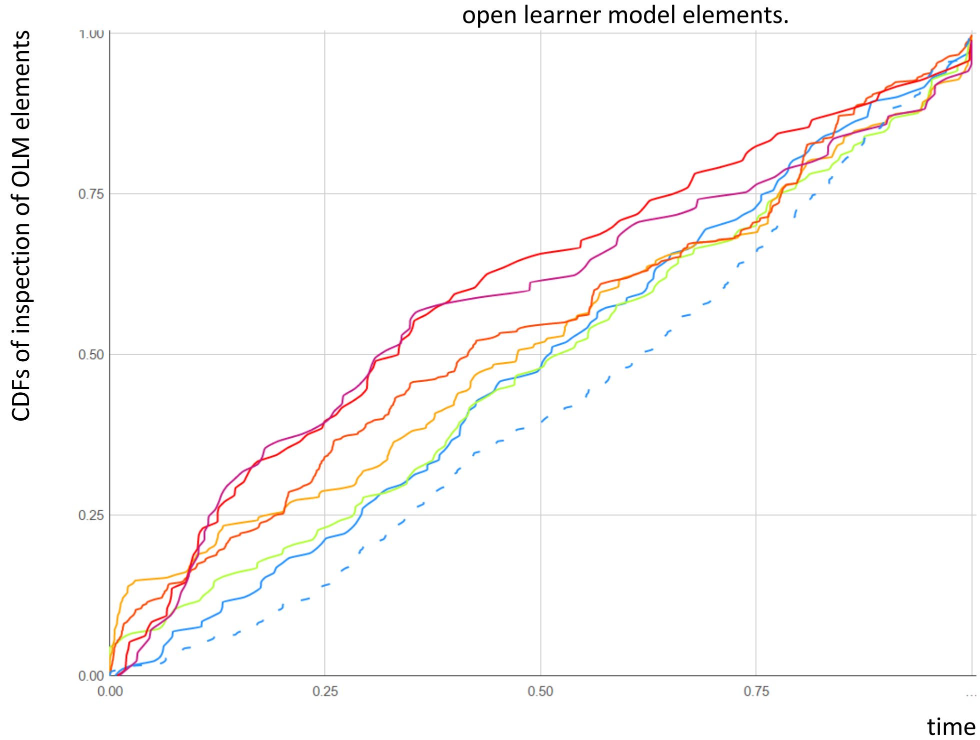


Key:
- **Dotted Blue** – question is presented
- **Blue** - answering questions
- **Green** - inspection of the model – overall
- **Yellow** – inspection of the model – topic level
- **Orange** – inspection of the model – concept level
- **Red** – inspection of the model – belief level
- **Purple** – inspection of domain model content.

Figure 12: MA(I) group cumulative distribution frequencies of interface usage across time.

BO(I) Group: CDFs of inspection of OLM elements (15 Participants)

Overall: linear exponent for interaction. Positive exponent for the inspection of finer granularity open learner model elements.



Key:
- **Dotted Blue** – question is presented
- **Blue** - answering questions
- **Green** - inspection of the model – overall
- **Yellow** – inspection of the model – topic level
- **Orange** – inspection of the model – concept level
- **Red** – inspection of the model – belief level
- **Purple** – inspection of domain model content.

Figure 13: BO(I) group cumulative distribution frequencies of interface usage across time.

7.1.3 Is drill down always used when inspecting the open learner model?

Both of the above measures give information on the frequency of use of components at each level of granularity. Table 16 shows that the majority of inspections of the OLM go beyond initial summary level of information to more detailed elements of the model. The SDs indicate there is a lot of variation within the dataset, potentially highlighting students' individual strategies and contextual use. The mathematics groups follow a normal distribution, with drill down to concept level the most common on average. The average distributions are also similar in the two groups where the model already contained a good level of information at the point of first access.

Table 16: frequencies of deepest level levels of OLM inspection.

Mean number of sequence occurrences	Summary	Topic	Concept	Beliefs[6]	Domain
MA (U)	4.06 (SD 4.23)	5.13 (SD 5.14)	7.06 (SD 8.00)	5.44 (SD 4.13)	4.38 (SD 3.12)
BO (U)	2.54 (SD 4.01)	1.85 (SD 2.23)	1.23 (SD 3.03)	1.85 (SD 1.99)	2.23 (SD 2.55)
MA (I)	0.81 (SD 1.76)	2.13 (SD 4.92)	5.38 (SD 8.17)	2.94 (SD 2.59)	2.81 (SD 2.46)
BO (I)	0.67 (SD 1.11)	1.20 (SD 1.97)	3.73 (SD 5.69)	2.93 (SD 3.45)	2.47 (SD 2.61)

Table 17 highlights a consistency between the 4 experimental groups in terms of the finite number of navigational possibilities that participants could take (see Section 5.6.4 for the underlying state model). The probability (as inferred from the proportion of student actions) of a drill down to some depth is high in the instantiated groups and the maths domain (MA (U) $p=0.8$, MA (I) $p=0.92$, BO (I) $p=0.91$) and slightly less for the Boolean group with an initially empty model ($p=0.62$). The most common course of action upon inspection of the model in all cases is to drill down to the depth of *concept* level, and the high probability of inspecting a

[6] N.B. beliefs and domain are to the same level of granularity; sequences will be counted twice if the participant has accessed both of these resources.

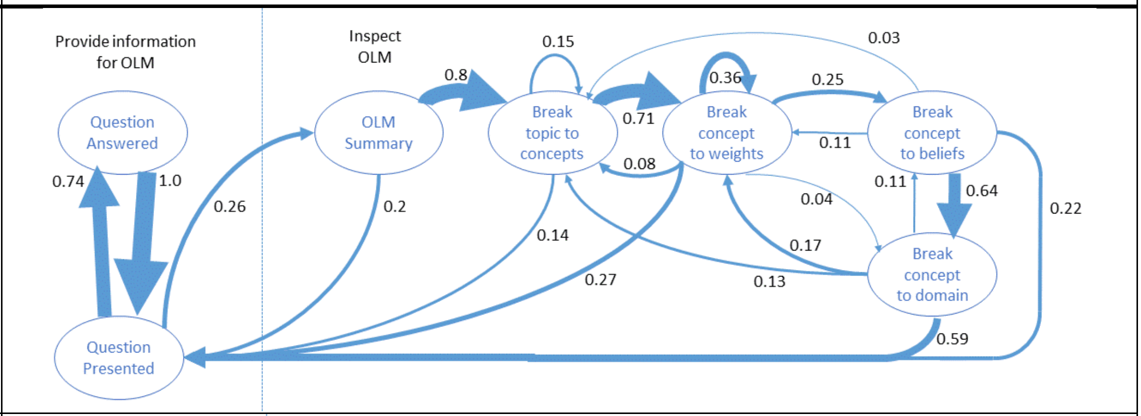
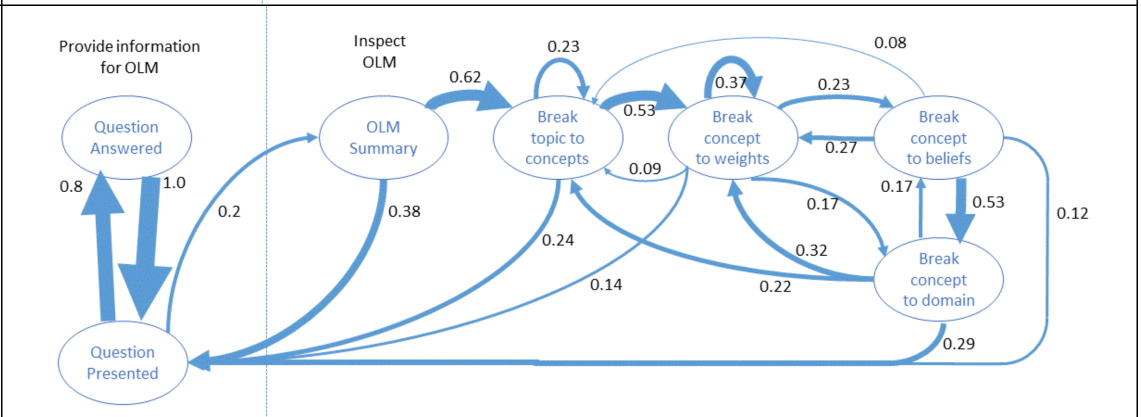
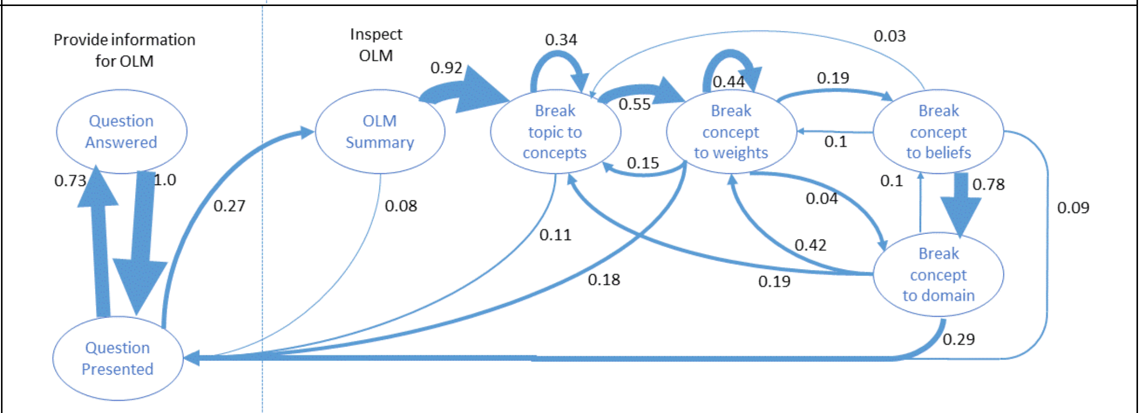
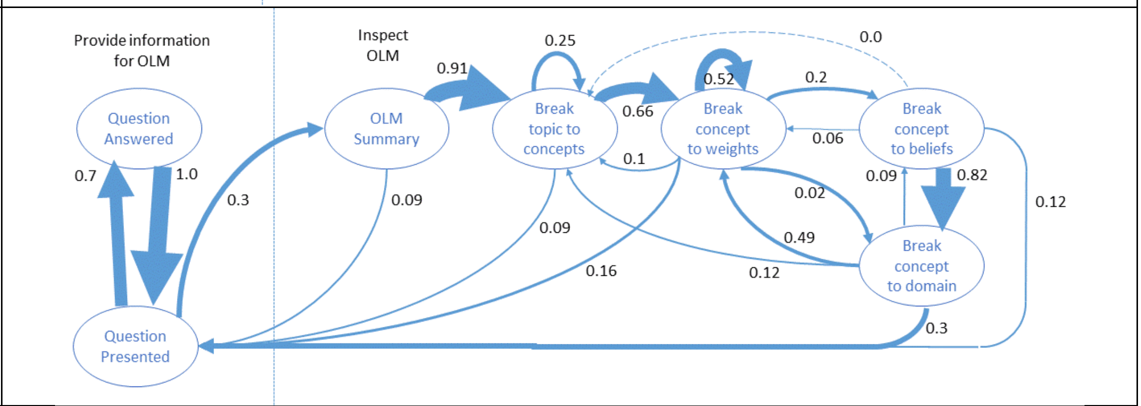
further concept following this would indicate the behaviour of skipping through the concepts to get an overview at this depth of inspection. Approximately p=0.2 of inspections, in each group, were from the concept depth to the belief depth of drill down, showing a consistency, regardless of domain or of level of information in the model.

In the groups where the model is initially empty, breaking from a shallow level inspection to answer questions is more likely (MA (U) p=0.2, BO (U) p=0.38) than when the model is initialised (MA (I) p=0.08 and BO (I) p=0.09). This confirms that when the model contains a greater volume of information, students do drill down further into it.

In each of the four groups, the most common action after inspection of a belief in the learner model is to inspect the domain model, and this is more likely when the model is initialised (MA (U) p=0.64, BO (U) p=0.53, MA (I) p=0.78, BO (I) p=0.82), thus showing that the beliefs from the OLM and the domain model representation for the same content are accessed together on more occasions than not. The converse is not observed for accessing domain content first, even though this information is at the same level of granularity.

Paths back from deeper levels of drill down are utilised during an inspection episode within the learner model, showing that students are on occasion returning to a coarser level of granularity and inspecting a further informational focus. The probability of breaking away from the cycle of questioning to inspect the learner model is between p=0.2 and p=0.3 across the four experimental groups.

Table 17: state models for system interaction showing average probabilities of transition.

Group	State Model
MA U	
BO U	
MA I	
BO I	

*Numbers indicate the probability that of transitioning from one state to the next, based on actual student use.

7.1.4 Summary

The evaluation reports that students indicated their acceptance of the individual granular representations of the OLM and the approach to using a drill down structure to organise these, with the majority stating that both are understood and useful. Overall, there is a good level of interaction with all aspects of the technology upon which to base analysis. Participants made the decision to drill down into the model in some form on the majority of inspections of the OLM, and the behavioural description presented here suggests that there are conditional elements as to when to make use of drill down, and that this behaviour is fairly consistent across time. There is a notable variation with regard to greater use of coarser grained aspects of the model (as would be consistent with overviewing its contents) during periods where there are larger volumes of information presented or there are many changes to initially observe. There is also evidence of participants cycling through viewing information in the model at 'concept' level which supports this. For the majority of interaction students also tended to answer questions in very short sharp bursts, with a strong bias towards 1-2 questions at a time in all experimental groups. Key observations are summarised in Table 18.

Table 18: how is an OLM with a drill down approach used? - key observations.[7]

1(a) Is an open learner model with a drill down approach accepted by its users?	**Yes.** The technology had a good level of use and the majority of participants indicated that with reference to each level of granularity and the drill down approach that it was understood and useful. Acceptance measures are limited to participants who chose to engage with the study only.
1(b) Is use of a drill down approach in the open learner model consistent across time?	**Yes.** Use of the drill down structure was consistent across the studies with reference to conditional deep drilling of the model. There is also some evidence pointing towards using the coarse grained elements for overview when there is more information or a greater number of updates.
1(c) Is drill down always used when inspecting the open learner model?	**No.** The majority of inspections of the OLM made use of drill down in some form, however the depth and extent to which it was drilled varied, indicating that there are conditional aspects to this.

[7] No statistical significances are included at this stage, as user acceptance and level of use is subjective, and in the context of the research question this does not have a benchmark. Information included in this chapter is sufficient to confirm that users' interactions satisfy the baseline criteria for being able investigate metacognitive effect and cognitive context, as part of the research.

Is students' use of drill down consistent across engineering domains?	**Yes.** General acceptance and general usage patterns are consistent across the experimental groups, including conditional deep drilling of the OLM. **No.** However, there are some variances in the extent of use and a slightly lower user acceptance for the Boolean logic groups, potentially relating to differences in student familiarity or method of engagement with the domain content.
Is students' use of drill down consistent when there are a large number of updates, such as a data import?	**Yes.** The frequency with which the learner model is updated is similar. There is evidence of both conditional deep drilling and overviewing of the model. **No.** There is an initial deeper level of inspection and slightly greater agreement that the model is understood and useful. This is potentially so as there were a greater number of changes to have initially observed.
Do students make use of open domain model information in the same way as OLM information, with a drill down approach?	**Yes.** Inspection of open domain model content is used with the same frequency as the learner's own beliefs and these are often used together in conjunction, potentially as a method through which interpretation takes place. **No**. Domain information is rated as slightly more readily understood and useful than the learners' own beliefs at the same granularity. Learners accessed their own model information then immediately afterwards the domain model.

7.2 What is the Impact of an Open Learner Model with a Drill Down Approach on Self-Assessment Accuracy?

Analysis considered here contributes towards an evaluation of how an OLM with a drill down approach impacts on learners' ability to self-assess. This mandates a consideration of the change student accuracy of completing this action across time and also of whether this is to a greater or lesser extent when users engage with drill down. Identification of the points of measurement for the comparison across time is included in Appendix 5.

7.2.1 Will student self-assessment accuracy increase over time when using an open learner model with a drill down approach?

Table 19 shows the mean discrepancy between the student behavioural model and diagnostic model state at each quarter of interaction and suggests <u>that there is a slight reduction in average discrepancy in the case of 3 of the groups, between the initial and final state. Changes are not statistically significant</u>[8]. There is a level of variance with some students increasing in discrepancy and some managing to reduce this, but this is not the case for all participants.

[8] Two tailed T-test.

Table 19: mean discrepancy between the behavioural and diagnostic models, per quarter.

Group	Initial	End of quarter 1	2	3	4	Difference between first and last measurement
MA U	No data available	0.26 (SD 0.17)	0.24 (SD 0.12)	0.24 (SD 0.16)	0.21 (SD 0.16)	-0.05 (SD 0.17)
BO U		0.26 (SD 0.24)	0.25 (SD 0.18)	0.21 (SD 0.16)	0.20 (SD 0.13)	-0.07 (SD 0.28)
MA I	0.10 (SD 0.10)	0.10 (SD 0.10)	0.09 (SD 0.09)	0.12 (SD 0.11)	0.11 (SD 0.12)	0.1 (SD 0.13)
BO I	0.20 (SD 0.15)	0.18 (SD 0.12)	0.15 (SD 0.08)	0.16 (SD 0.08)	0.16 (SD 0.11)	-0.04 (SD 0.14)

A comparison of ability to increase in self-assessment accuracy during interaction against students' own indication of their self-assessment abilities shows a significant[9] correlation (Table 20). This suggests that students giving an indication of their self-assessment ability may be used as a proxy measure to identify those in the sample who are likely to effect greater increases in accuracy on their self-assessments whilst using an OLM that implements a drill down structure in its interface design. (Supporting analysis is included in Appendix 6, Section 1.)

Table 20: students' indication of their self-assessment ability correlated with changes in self-assessment during interaction.

Experimental Group	Correlation between students' indication of their self-assessment ability and their ability to increase the accuracy of their self-assessment during interaction. Description	R	Significance
Ma(U)	Indicates stronger self-assessor, greater increases in accuracy	0.6335	**p<0.05**
Ma(I)	Indicates stronger self-assessor, greater increases in accuracy	0.5772	**p<0.05**
Bo(U)	Indicates stronger self-assessor, greater increases in accuracy	0.6567	**p<0.01**
Bo(I)	Indicates stronger self-assessor, greater increases in accuracy	0.4637	**p<0.05**

As included in Table 21, interaction data shows that it is significant[10] (p<0.001) in the MA(I) group that those who accessed the OLM more frequently made bigger increases in self-assessment accuracy over the time period of one hour. In each of the three other experimental groups the frequency of use of the OLM was not related to changes in self-assessment accuracy. (Supporting analysis is included in Appendix 6, Section 2.)

[9] Spearman rank correlation

[10] Spearman rank correlation

Table 21: students' frequency of use of the OLM correlated with changes in self-assessment during interaction.

Experimental Group	Correlation between frequency of OLM access and student ability to increase the accuracy of their self-assessment during interaction.		
	Description	R	Significance
Ma(U)	More frequent OLM use, greater increases in accuracy	0.6076	**p<0.001**
Ma(I)	More frequent OLM use, greater decrease in accuracy	-0.0022	Not significant
Bo(U)	More frequent OLM use, greater decrease in accuracy	-0.3741	Not significant
Bo(I)	More frequent OLM use, greater decrease in accuracy	0.2330	Not significant

7.2.2 Will student self-assessment accuracy increase to a greater extent when greater use is made of drill down?

Table 22 compares changes in student self-assessment accuracy against their use of each level of drill down and indicates that the two groups who used the system in the Boolean logic domain decreased in self-assessment accuracy with more frequent access of beliefs and overview levels of the OLM (significant for uninstantiated for topic level ($p<0.01$), belief level ($p<0.01$) and domain information ($p<0.05$)). In contrast, the two groups who used the system in the mathematics domain showed a non-significant increase in accuracy with more frequent use of drill down. Overall this gives an inconclusive finding and may additionally may suggest that learners were struggling to work with and interpret the OLM content in the context of the Boolean logic domain, or that there is an issue of correlation versus causation: for example, frequency of OLM use may be influenced by other factors. (Supporting analysis is included in Appendix 6, Section 3.)

Table 22: students' frequency of use of drill down correlated with changes in self-assessment during interaction.

Experimental Group	Level of Drill Down (Spearman Rank Correlation)			
	Topic	Concept	Belief	Domain
Ma(U)	R=0.3642 not sig	R=0.0300 not sig	R=0.1341 not sig	R=0.2529 not sig
Bo(U)	**R=-0.7195 p<0.01**	R=-0.2118 not sig	**R=-0.2118 p<0.01**	**R=-0.6322 p<0.05**
Ma(I)	R=0.1791 not sig	R=0.1568 not sig	R=0.1048 not sig	R=0.0842 not sig
Bo(I)	R=-0.2502 not sig	R=-0.2213 not sig	R=-0.2161 not sig	R=-0.1774 not sig

7.2.3 Summary

Across each of the 4 experimental groups there are no significant changes in student self-assessment accuracy during interaction for all participants, although some participants did manage to achieve increases in accuracy. It is possible to use student stated indication of their own self-assessment ability as a proxy measure to identify those who will be able to increase the accuracy of the self-assessments to a greater extent. In one experimental group more frequent access of the OLM correlated with greater increases in self-assessment accuracy, however this is inconclusive for the other experimental groups. It is also inconclusive as to whether students will increase in self-assessment accuracy to a greater extent when greater use is made of any given levels of drill down. Key observations are summarised in Table 22.

Table 23: changes in self-assessment accuracy - key observations.

2(a) Will student self-assessment accuracy increase over time when using an open learner model with a drill down approach?	**No.** Overall there was a slight non-significant increase in accuracy in 3 of the 4 groups but this did not generalise across participants. Students indicating their self-assessment ability can be used as a proxy measure to identify those who improve to a greater extent.
2(b) Will student self-assessment accuracy increase to a greater extent when greater use is made of drill down?	**No.** Findings are inconclusive as to the extent of a relationship between frequency of use of drill down and self-assessment accuracy increases. More frequent access did correlate with greater increases in accuracy for one experimental group.
Is students' use of drill down consistent across engineering domains?	**Yes.** It is consistent that no significant increases in self-assessment accuracy are observed, and it is inconclusive in both domains whether frequency of access relates to changes in self-assessment.
Is students' use of drill down consistent when there are a large number of updates, such as a data import?	**Yes.** Withstanding the differences in the precision of the data from different levels of model completeness, it is consistent that there are no significant changes in self-assessment accuracy during interaction or with reference to extent of drill down use.
Do students make use of open domain model information in the same way as OLM information, with a drill down approach?	**Yes.** The use of the drill down information for domain inspection is consistent with the inspection of beliefs.

7.3 What is the impact of an open learner model with a drill down approach on the support for the regulation of cognition?

The thesis now considers participants' use of the OLM technology in terms of interaction behaviour that might support tasks that are important to the regulation of cognition. This section focuses on students' inspections of their problems and use of the technology in a focused way as might indicate behaviour consistent with actions such as problem solving.

7.3.1 Will students use drill down in the OLM to inspect information about problems?

Section 5.6.7.1 describes the calculation for determining the difference between the behavioural model state for an item selected for drill down in the OLM and the average state of the behavioural model, at the point at which the navigation occurred. Figure 14 shows that in both groups where the model was initially empty, and both groups in the maths domain, similarities are present that indicate that <u>for topic level inspections the items selected are about the average state of the model, whereas for concepts this is less than the average state, and for beliefs (and corresponding domain content), the average of this is less again (significant[11], MA(U) p<0.01, (BO(U) p<0.05).</u> This pattern is also observed to a slightly lesser extent in the model that was initialised in the maths domain (significant[12] MA(I) p<0.01). Among inspection of information at each level of drill down, the selection of items weaker than the average state of the model is only significant for beliefs for mathematics domain when the model is initially empty (p<0.05). While in the other groups this is an interesting

[11] significances between mean state and state inspected calculated with a two-tailed T-test.
[12] significances between mean state and state inspected calculated with a two-tailed T-test.

characteristic of the data, it potentially requires further work to establish a more definite answer. Data is included in Appendix 6, Section 4.

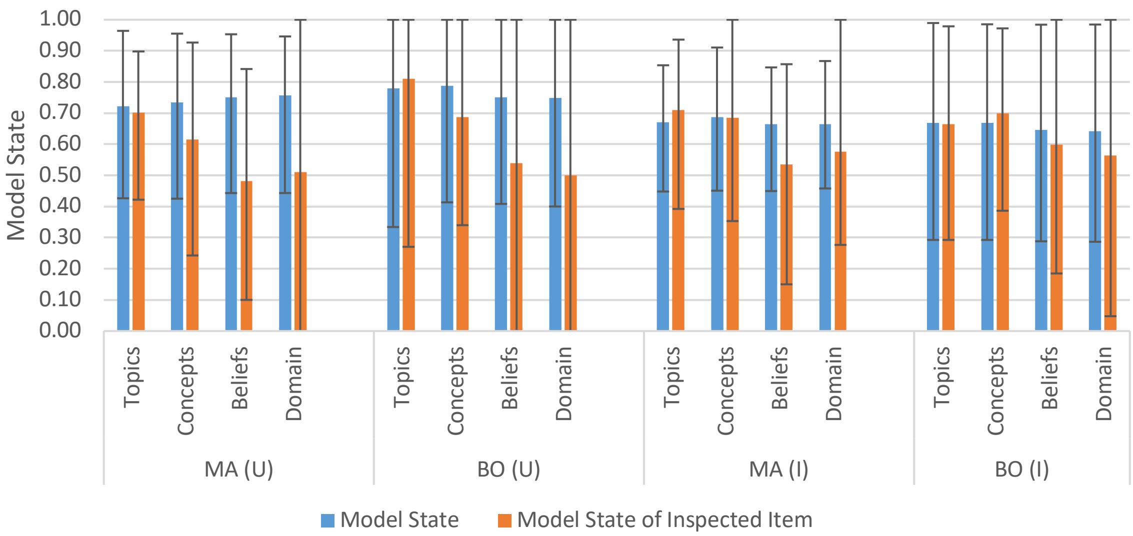
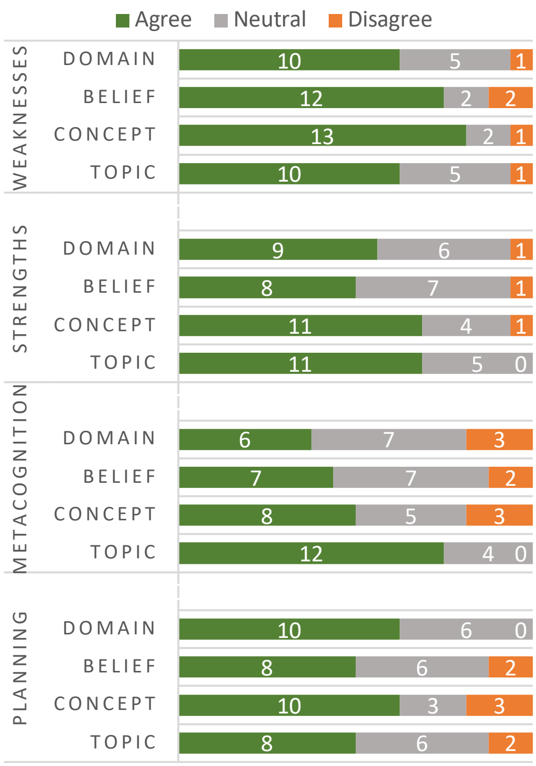
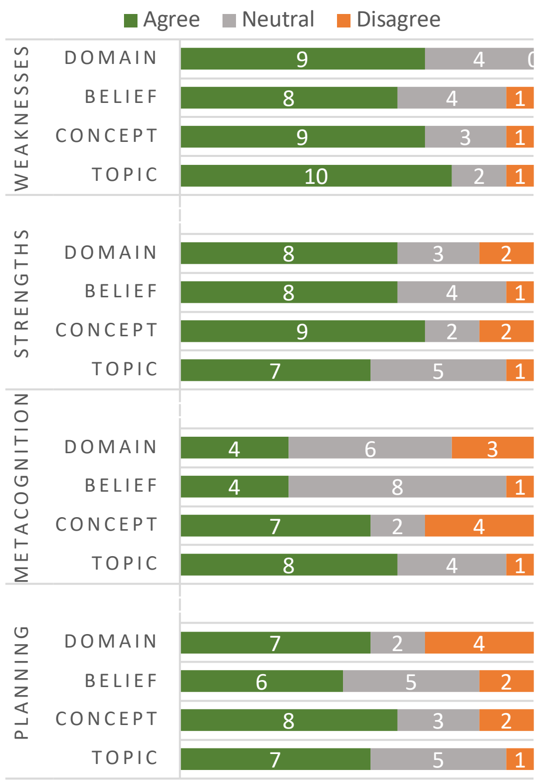


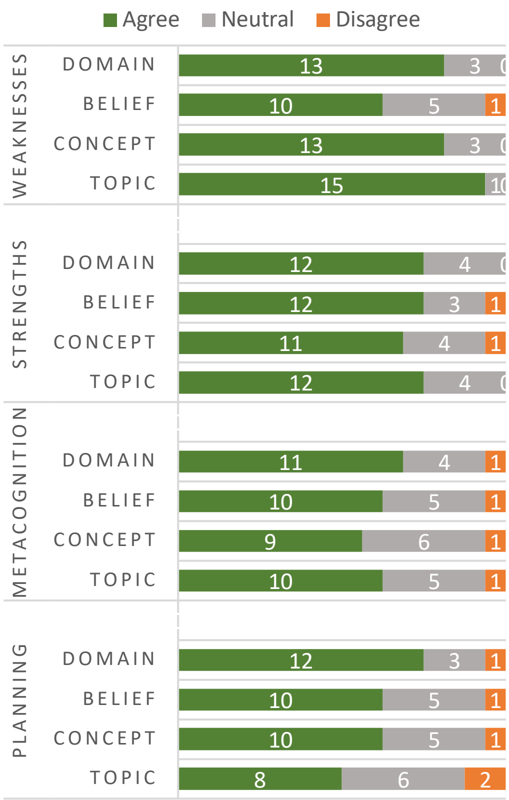
Figure 14: behavioural model state and deviation from the model state at point of access.

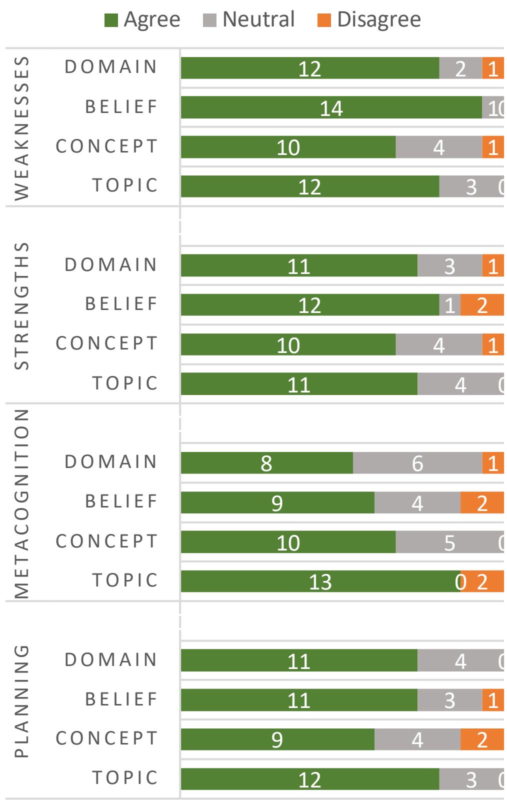
In the post usage survey (Table 24), participants were asked for their perceptions of reasons for use, relevant to aspects of actions important to the regulation of cognition (see Section 5.5.2 for commentary and Appendix 3 for the survey). In each of the four groups, while there is some variance in student responses, the OLM with a drill down approach was thought by the majority of students to be useful for identifying weaknesses and areas of improvement on each level of granularity. This was consistent across each of the experimental groups. This is likewise the case for identifying strengths. Although in each case it is not useful to all participants for these purposes, as some students stated neutrality on the matter and one or two participants disagreed; the majority used it for these purposes. Relevant to actioning the information viewed at each level of drill down students were asked whether they were using the information for planning subsequent actions. Approximately half the participants agreed that they had used the technology for some aspect of this, with the

Table 24: survey responses - reasons for opening the learner model.

<table>
<tr><td></td><td>Maths</td><td>Boolean</td></tr>
<tr><td>Uninstantiated</td><td>


</td><td>


</td></tr>
<tr><td>Instantiated</td><td>


</td><td>


</td></tr>
</table>

majority of the other participants giving a neutral response. In each of the four groups participants stated similar usage across each of the four grain sizes considered. Students were also asked whether their use of the visualisations helped them think about how they were approaching their learning of the subject matter (labelled as 'metacognition' in Table 24). Over half the participants agreed that they had used the model for this in the instantiated groups, and in the uninstantiated groups over half agreed that they used the coarse grained parts of the technology for this – there was greater neutrality for the finer grained visualisation components (consistent across both domains).

7.3.2 Will students use drill down in the OLM to inspect information about areas of uncertainty?

Section 5.6.7.2 describes the calculation for determining the difference between the diagnostic model state for an item selected for drill down in the OLM and the average state of the diagnostic model, at the point at which the navigation occurred. Figure 15 shows that while some variance exists between participants, students' use of the drill down approach averages out close to the average state of the model. This is consistent for all levels of drill down, for each of the 4 groups. There are no significant deviations from this, as in the case for the inspection of behavioural model information (Section 7.3.1). Data is included in Appendix 6, Section 5.

Of the slight deviations in the data, in the case of the two mathematics groups, the average for inspected items is slightly greater than the average model state. Of the two instantiated model groups, the state of the diagnostic model for items inspected is consistent across all levels of drill down. For the two groups where the model was initially empty the more detailed components of the drill down are accessed for concepts where the learner is slightly more confident for mathematics, and slightly less confident in the case of Boolean logic.

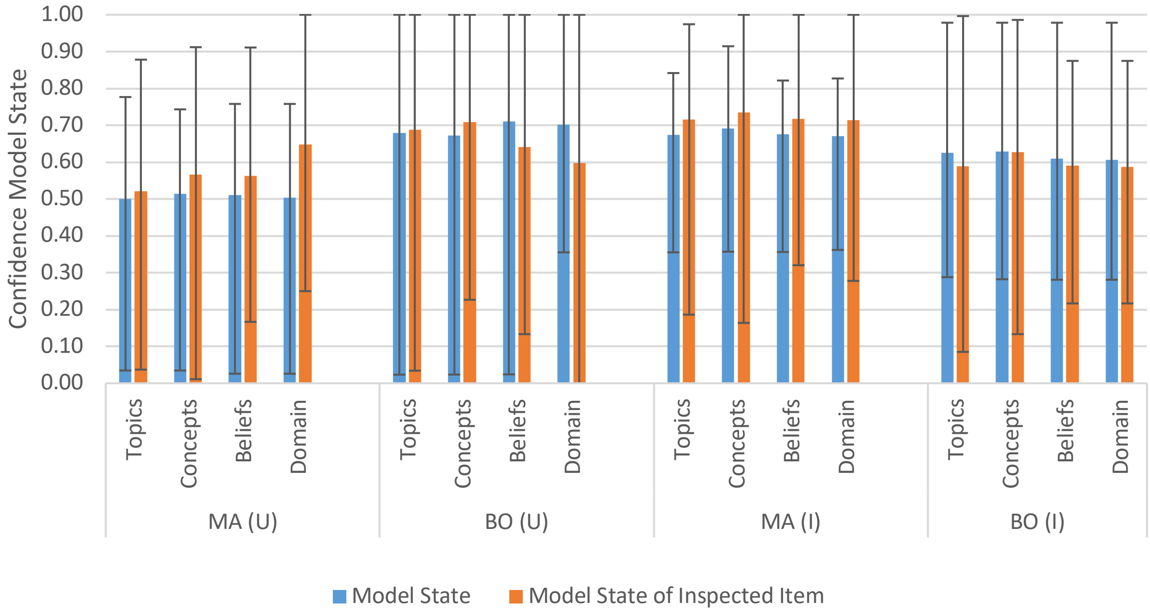


Figure 15: diagnostic model state and deviation from the model state at the point of access.

7.3.3 Will students use an OLM with a drill down approach to focus on one domain area at a time?

As per the calculation outlined in Section 5.6.7.3 , Figure 16 shows the average number of question blocks ago that the same domain concept was encountered for updating (i.e. how recently has the domain content been updated). While there is some difference in terms of the mean averages at different levels of drill down, the median values lie much closer to 1.0, indicating that the learner was working with a question on this concept in the last round of updating the model. There is a strong bias towards the domain content inspected coming from the last 1 and in some cases 2 rounds of questioning. This is consistent across different levels of drill down and experimental groups.

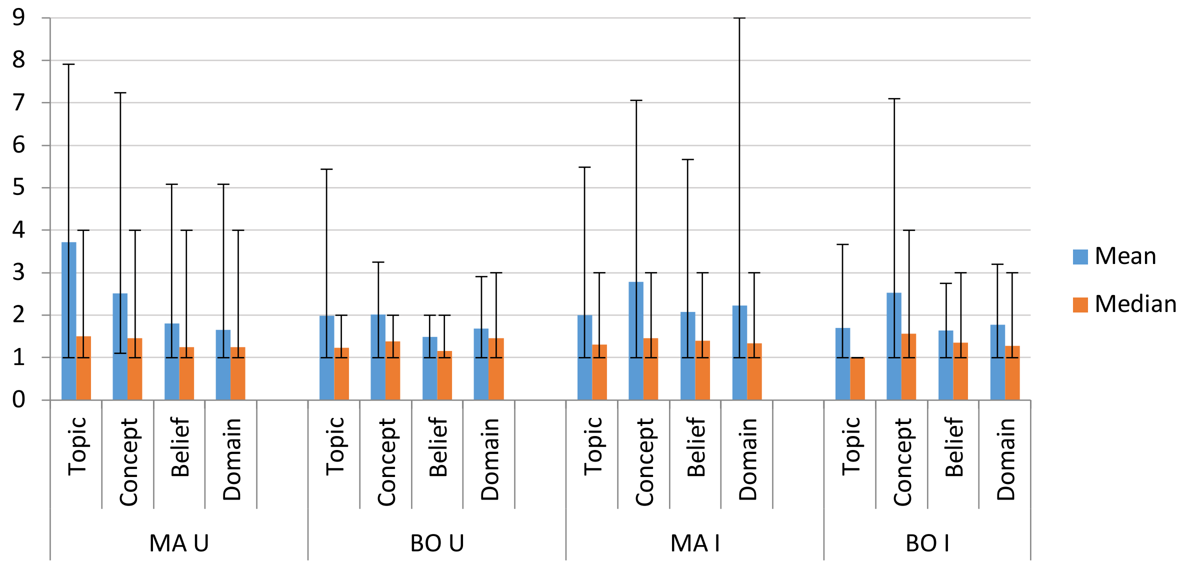


Figure 16: for parts of the model inspected in more detail, how many question-answering sessions ago were the same topic/concept encountered?

With the mathematics group with the initially empty model, there is significant suggestion that the deeper the model is drilled the more recently the concept was worked with in answering questions (significant[13] MA(U) $p<0.001$), but it is not significant in other groups. These results may, for example, indicate that students are seeking detailed feedback on items they are currently focusing on, whereas higher level inspection may be to view general progress. (Supplementary data is included in Appendix 6, Section 6.)

On the post usage survey (Appendix 3 and Figure 17), participants were asked to indicate reasons for viewing their learner model in the context of providing information. In each of the four groups the majority of participants agreed that they inspected the OLM to see general progress, and overall there was strong disagreement with the statement stating that they weren't sure why they were inspecting the OLM, indicating that students had a purpose in mind. With reference to the informational content, students agreed in around half of cases

[13] significances between different drill down levels calculated with a 1 way ANOVA (related)

that this was for information about the last question answered, or a previous question (although there is some variance between groups), and mostly disagreed with the idea that they were inspecting the model with reference to a future update (although there were some who indicated they were likely to use it for this purpose).

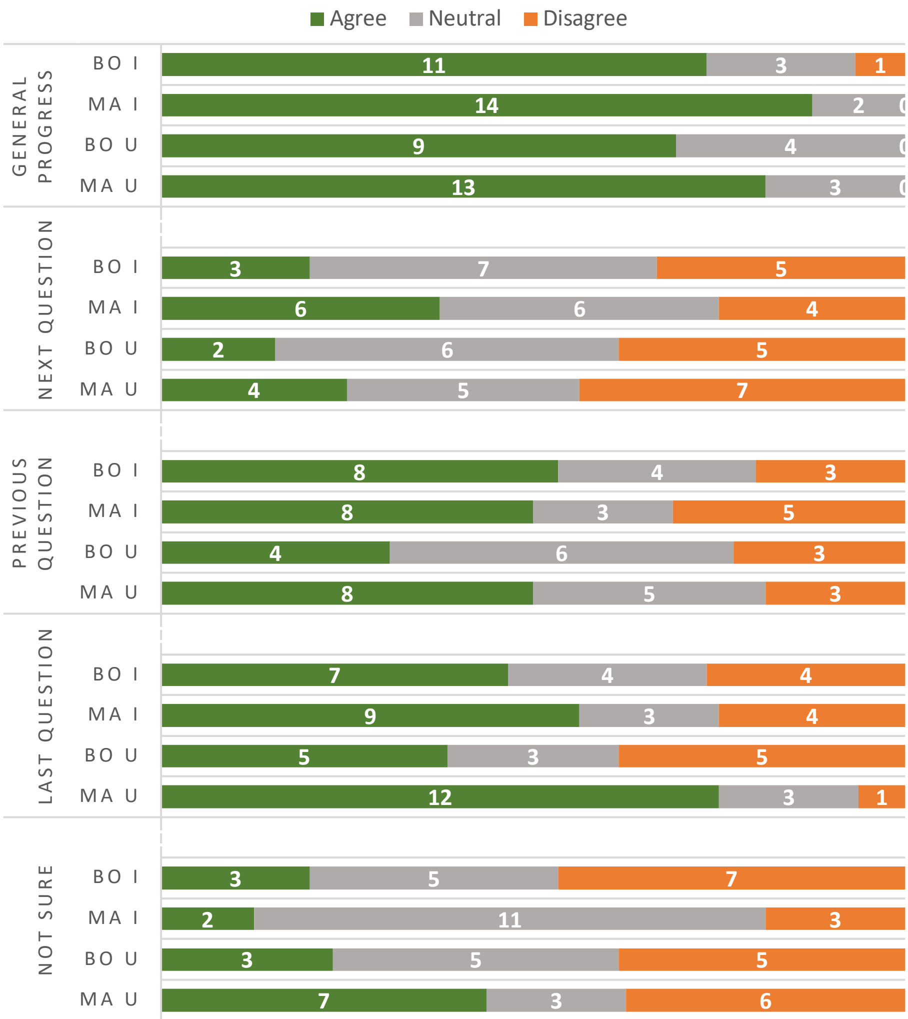


Figure 17: stated reasons for viewing the learner model, in the context of updating it.

Overall this supports the usage data presented in this section. <u>While there is a strong bias to inspecting information that is very recently updated in the learner model, there is also the desire to see general process</u>, which would explain the variance in the data. Findings are fairly consistent across levels of drill down.

7.3.4 Summary

In the context of looking for supporting evidence for behaviour consistent with tasks regulatory to cognition the thesis has reported on: (i) students' use of the OLM to identify problems; as (ii) distinct from areas of uncertainty; and (iii) evidence of focussed interaction that would be consistent with tasks such as problem solving. This section of the analysis has confirmed that students exhibited much behaviour consistent with these tasks as is evidenced through their use of drill down. The analysis has reported that the behavioural model state of coarse level inspections is about the model average, that students made use of drill down for identified problems, and that students drilled down deeper when problems were of a greater magnitude. In the post usage survey students further confirmed that they were using the OLM to identify problems in addition to their strengths (although these were less often inspected in detail) and some learners used the OLM for elements for planning subsequent actions or a consideration of how they're approach their learning. This careful and detailed inspection of student problems is distinct from the use of drill down to find information in domain areas in which they are less confident, where no correlation is observed. There was also an indication in one of the experimental groups that the more recently a concept was updated, the deeper the model was drilled, which is the use of drill down for detailed inspection of domain areas in current focus. Combined with the observation of Section 7.1 that students mostly worked with short questions bursts of 1-2 questions of a time, this suggests a potential cycle of answering a question and immediate viewing of feedback, with a conditional use of drill down if the OLM identifies problems in the student's behavioural model. In the post usage survey students often indicated that their reasons for switching from questioning to view the OLM was to see information on the last/previous question and to a greater extent to also observe general progress. There was disagreement with the idea that students did not know why they were viewing the OLM, which indicates that they perceived that their use of it was purposeful. Key observations are included in Table 25.

Table 25: intervention usage compared to the behavioural model - key observations.

3(a) Will students use drill down in the OLM to inspect information about their problems?	**Yes.** Students made use of drill down to explore problems in greater detail and drill down deeper when problems were of greater magnitude. This observed in both groups in the maths domain ($p<0.01$, $p<0.01$) and where the model is initially empty in the case of Boolean logic ($p<0.05$).
3(b) Will students use drill down in the OLM to inspect information about areas of uncertainty?	**No.** There is no significant pattern in the data to suggest that learners' confidence in their abilities relates to their decision to drill into the OLM.
3(c) Will students use an OLM with a drill down approach to focus on one domain area at a time?	**Yes.** There is a bias towards domain content being selected for drill down that relates to the last update of the model (reported in this section), and that the model is updated in short sharp bursts of 1-2 questions at a time (reported in Section 7.1). The post usage survey confirms that over half of the participants were consciously desiring feedback on a question just answered, and the majority also indicated interest in general progress.
Is students' use of drill down consistent across engineering domains?	**Yes.** In both domains students drilled deeper when the model identified problems and drilled deeper when problems were greater. In both domains this is distinct from students' use of drill down with relation to their confidence in their abilities, which showed as not related. Likewise, in both domains there is a strong bias towards viewing information related to the last domain area updated, and (from Section 7.1) the majority of questions being in short bursts.
Is students' use of drill down consistent when there are a large number of updates, such as a data import?	**Yes.** While the dataset is slightly noisier when the model is instantiated at the point of first viewing, it is consistent that students primarily have an interest in areas that are problematic, that they perform careful inspection of the content and answer further questions on these elements. This is not related to student confidence and is with a focused area of the domain, of which subsequent attempts are made to update the behavioural model in the same area.
Do students make use of open domain model information in the same way as OLM information, with a drill down approach?	**Yes.** There is an element of consistency shown between the inspection of domain information and beliefs in the learner model. Section 7.1 suggests that these are often used together, and so the conditions for inspection of the domain are again where problems are existing distinct from lack of confidence in the area, and with domain content recently updated. The pattern of short updates followed by immediate inspection means it may also feed forward to the next question attempt.

7.4 Comparison of Experimental Groups

This chapter has reported on 4 experimental groups to highlight elements of consistency across domains (basic engineering mathematics, and Boolean logic) and across the case where the model is already complete on the occasion that students first access the OLM.

7.4.1.1 Intervention Use Across Engineering Domains

The summaries presented throughout this chapter that refer to comparisons between domains are collated in Table 26. This chapter has confirmed that there is similarity in terms of acceptance and usage, and conditional deep drilling over the model combined with potential overviewing behaviour. However, overall there was a lesser level of interaction and acceptance in the Boolean logic domain than for mathematics. In the Boolean logic domain use of drill down started slightly later in interaction, and as stated in Section 7.1 there is potential that students may have had greater difficulty engaging with the Boolean logic content than the mathematics. This is potentially because of different levels of familiarity with the domain content as Boolean logic is slightly less often directly encountered in participants' studies than basic engineering mathematics, or because other interaction strategies may have been applied by students during the evaluation.

Table 26: Is students' use of drill down consistent across engineering domains?

Section 7.2 (acceptance and use)	**Yes.** General acceptance and general usage patterns are consistent across the experimental groups, including conditional deep drilling of the OLM. **No.** However, there are some variances in the extent of use and a slightly lower user acceptance for the Boolean logic groups, potentially relating to differences in student familiarity or method of engagement with the domain content.
Section 7.3 (self-assessment accuracy)	**Yes.** It is consistent that no significant increases in self-assessment accuracy are observed, and it is inconclusive in both domains whether frequency of access relates to changes in self-assessment.
Section 7.4 (supporting regulation of cognition)	**Yes.** In both domains students drilled deeper when the behavioural model identified problems and drilled deeper when problems were greater. In both domains this is distinct from students' use of drill down with relation to their confidence in their abilities, which showed as not related. Likewise, in both domains there is a strong bias towards viewing information related to the last domain area updated, and (from Section 7.1) the majority of questions are in short bursts.

It is consistent that both domains showed no significant change in self-assessment accuracy and that a proxy measure of students indicating their self-assessment ability can be used in both domains to identify those who are likely to increase in self-assessment accuracy to a greater extent. Likewise, there is evidence in both domains of drilling down into the model and performing detailed inspections of problems, and in both domains use of drill down was not related to student confidence in their ability to correctly apply domain knowledge. In both domains learners also worked with short bursts of questions followed by immediate use of drill down in the same domain area, indicating an element of focus. Each of these elements of interaction are generalisable across the two engineering domains considered.

7.4.1.2 Intervention Usage with Different Volumes of Information

To allow investigation of student interaction with a complete OLM and in part simulating the circumstance where a large import of data from another system (e.g. CMS) may have taken place, two of the experimental groups were only granted access to the OLM once it was mostly complete. The summaries of this chapter are collated in Table 27 and confirm that beyond the initial viewing period when access to the model is granted, there is great similarity with use of the model as compared to when initially empty (although there is less interaction data to evaluate in the instantiated groups). Across both conditions students showed evidence of conditional deep drilling of the model in areas that are identified as problematic, and then continued to update the model using short bursts of questions and inspect the same area of the domain in detail following this. This might suggest a baseline interaction strategy that is generalisable. Again, whether instantiated or initially empty, there is no significant change in self-assessment accuracy during interaction, but students indicating their self-assessment ability is an accurate proxy to identify those who would increase in accuracy to a greater extent. This potential baseline also shows interaction behaviour consistent to support the argument that the OLM and drill down are being used to support elements of tasks that are regulatory to cognition.

As might be expected in the instantiated groups, there is an initial surge in the use of the drill down, as there is much to inspect once access is granted to the OLM. In the post-usage survey the majority of participants also indicated that they used to OLM for helping plan subsequent actions and thinking about how they're approaching their learning. Agreement with these statements is stronger than when the model is initially empty. Together this gives weight to the argument that students may be using this period of interaction to identify and plan which areas of the domain need attention, as is further evidenced by their subsequent engagement in domain areas that are much weaker than the average state of the behavioural model. Engagement with formative assessment in this way supports learners by providing them the opportunity to reflect on these domain aspects of their curriculum.

Table 27: Is students' use of drill down consistent when there are a large number of updates?

Section 7.2 (acceptance and use)	**Yes.** The frequency with which the learner model is updated is similar. There is evidence of both conditional deep drilling and overviewing of the model. **No.** There is an initial deeper level of inspection and slightly greater agreement that the model is understood and useful. This is potentially so as there were a greater number of changes to have initially observed.
Section 7.3 (self-assessment accuracy)	**Yes.** Withstanding the differences in the precision of the data from different levels of model completeness, it is consistent that there are no significant changes in self-assessment accuracy during interaction or with reference to extent of drill down use.
Section 7.4 (supporting regulation of cognition)	**Yes.** While the dataset is slightly noisier when the model is instantiated at the point of first viewing, it is consistent that students primarily have an interest in areas that are problematic, that they perform careful inspection of the content and answer further questions on these elements. This is not related to student confidence and is with a focused area of the domain, of which subsequent attempts are made to update the behavioural model in the same area.

7.4.1.3 Intervention Usage with Domain Content

While not explicitly part of the design of experimental groups, it is worthy of note in the context of Section 3.6 of the literature that contextual information is sometimes required for comprehension or problem solving, and to permit interpretation of detailed model aspects for the purpose of cognition. Observations relating to the open domain model content are summarised in Table 28, and the interaction data reported in this chapter suggests that students were willing and able to engage with this information and use it at points where detailed and systematic inspection of problems in the model is taking place. This is evident in students' activation of both domain model and learner model information in immediate succession, and in the survey which indicated that student beliefs were less readily understood than the coarser representations of the learner model. Such behaviour is indicative of an attempt at comprehension, potentially as a prerequisite to aspects of metacognitive activity.

Table 28: Do students make use of open domain model information in the same way as open learner model information, from the perspective of a drill down approach?

Section 7.2 (acceptance and use)	**Yes.** It is used with the same frequencies and often in conjunction with the learner's own beliefs, potentially as a method through which to interpret them. **No**. Domain information is rated as more slightly more readily understood and useful than learners' own beliefs at the same granularity. Learners accessed their own behavioural model information then immediately afterwards the domain model.
Section 7.3 (self-assessment accuracy)	**Yes.** The use of the drill down information for domain inspection is consistent with the inspection of beliefs.
Section 7.4 (supporting regulation of cognition)	**Yes.** There is an element of consistency shown between the inspection of domain information and beliefs in the learner model. Section 7.1 suggests that these are often used together, and so the conditions for inspection of the domain are again where problems are existing distinct from lack of confidence in the area, and with domain content recently updated. The pattern of short updates followed by immediate inspection means it may also feed forward to the next question attempt.

7.5 Discussion

The research question breaks down into a general consideration of how an open learner model is used by a learner, its impact on aspects of self-assessment and also its impact on support for metacognitive activity.

7.5.1 How is an open learner model with a drill down approach used?

Interaction has shown a fairly consistent use across time and different informational levels, but variances do indicate that students are using different strategies whilst engaging with the OLM, which should be expected as literature on individual differences suggests. The main notable aspect of interaction is conditional deep drilling of the model, particularly when the OLM showed problems in student understanding, and this is combined with evidence of overviewing behaviour, which is suggested by Boud and Molloy (2013) as actions consistent with a process of discovery that prerequisite to reflection. Students also verified the approach of using a drill down structure in an OLM is accepted in domains of basic engineering mathematics and Boolean logic, such that investigations with regard to metacognition and self-assessment have validity. There is fairly consistent usage across time and different information levels which also adds a level of validity to the findings that have generalised, and the baseline pattern of activity that has emerged. The baseline indicates that the majority of interaction is learners answering short bursts of questions (1-2 at a time) and then drilling into the model on the area that has just been updated, particularly when there are problems. This shows a level of focussed interaction, consistent with some tasks that are regulatory to cognition, with specific reference to learners' detailed inspection of problematic areas of the domain. Further question attempts are made on the same domain area which may potentially relate to students attempting to justify (to themselves or the system) or validate their current state of cognition. This is a pattern of behaviour that Hartman (1998) states may contribute to metacognitive development. Learners also stated that they have a particular interest in using the OLM to identify their problems, and also to identify strengths, general progress, and (in the case of some learners) aspects of planning and thinking about how they're approaching their learning. Each of these aspects are consistent with the established aims and reported use of open learner models, as summarised by Bull and Kay (2016).

7.5.2 What is the impact of an open learner model with a drill down approach on self-assessment accuracy?

There was no generalisable improvement or deterioration in self-assessment accuracy during interaction, but some students did increase in accuracy, and the proxy measure of asking students whether they are good at self-assessment has shown to be accurate at identifying which participants make improvements. Previous OLM studies (e.g. Mitrovic and Martin, 2007; Kerly et al., 2008) have shown generalisable increases in self-assessment accuracy over similar periods of interaction, however as suggested by Gabrielle (2003) this can be contextual. There is suggestion in one experimental group that more frequent use of drill down correlates with more improvements in self-assessment, potentially as part of the profile of a stronger learner or of the efficacy of a given interaction strategy. Even if no generalisable effect is seen over the period of the study, engagement with the technology inducts students in completing self-assessments and potentially raises the profile of the action, such that it may be applied subsequently to future learning episodes as it is contributory to metacognitive development (Desautel, 2009).

7.5.3 What is the impact of an open learner model with a drill down approach on the support for the regulation of cognition?

Tasks that are regulatory to cognition include problem solving, comprehension monitoring and evaluation, and it could be argued that students' interactions with the technology and inspection of their OLM is in part an act of comprehension monitoring. Students' regular inspections of the model and conditional use of the drill down means that they are completing this, even if this is not consciously or is as part of other actions. As Papies and Aarts (2016) highlight, students can often complete tasks regulatory to cognition without formally recognising them. Evidenced through students' navigation of the different elements of the technology, there is a baseline interaction pattern that is shown in this study. This pattern indicates a large proportion of student interactions are with short bursts of questions, with regular monitoring of areas of the domain immediately after their update, and that there is detailed inspection of problems when these are identified in the model. In their use of the technology students are participating in evaluation of the OLM's contents and minimally there is an indication that students were able to think critically to identify which areas of the OLM contained the greatest problems.

The different elements of this baseline activity together are consistent with the task of problem solving and strategies that may be associated with this regulatory process (see Mayer, 2013 and Section 2.1 for further behavioural descriptions). The evidence of focused interaction (that students are inspecting the model for the last domain item updated and working with one area of the domain at a time) is consistent with either a trial and error approach to updating the model or part of a hypothesis testing strategy (particularly when this is coupled with deep inspection problematic areas of the model). Likewise, the short question bursts indicate a regular inspection of feedback and close monitoring of changes in the model, following the process of updating. This could also indicate an element of systematic working and comprehension of formative assessment feedback, such that students are establishing their cognitive state and closely monitoring its update/change in a controlled way. This may also potentially be as part of a wider strategy to identify and/or rectify a problematic area of understanding as part of learning.

Additionally, students also indicated that their interest is for general information as well as for specific areas that are problematic, and there is evidence of overviewing behaviour in student interaction. This is potentially showing an element of synthesis with wider goals, strategies, and the identification of future foci or placement/contextualisation of the current domain focus within the wider context of what students are doing (potentially towards support for short term planning etc.). Both of these would be active enquiry processes that have the potential to encourage students to think about the state of their cognition through formative assessment, while undertaking appropriate navigational actions when interacting with the OLM. Overall, students indicated that they knew how they wanted to make use of the intervention, and this is an indicator of conscious and active engagement with the formative assessment content in a way that is consistent with actions that should support aspects of the regulation of cognition as part of a metacognitive process.

7.6 Summary

Evaluation has suggested that while students appear to be employing different strategies to interact with the model, a baseline pattern of interaction emerges that is consistent with problem solving strategies. This is relevant to suggesting that the technology, which is accepted by students, is being used in a way consistent with metacognitive actions. The presence of drill down is supportive of interaction of this type and it is appropriate to capture these aspects of interaction behaviour. The study has also shown that while students participated in self-assessment during use, there are no significant increases in self-assessment accuracy while using the technology. However, a proxy measure of students indicating their self-assessment ability can identify those who are more likely to increase in accuracy to a greater extent, potentially as part of a profile of a stronger learner. Study 2 will validate whether these usage patterns hold true over a longer period of student engagement, and whether active use alongside an ongoing course gives indication of the same types of interaction strategies.

Chapter 8

STUDY 2

This chapter reports on Study 2, according to the plan outlined in Chapter 5, and presents analysis and commentary on the sub-research questions as outlined in Section 5.1 . In contrast to Study 1, Study 2 focuses only on the basic engineering mathematics domain, is alongside an 8 week course, and has 27 participants. As in Study 1 this chapter first considers user acceptance of the technology of those who volunteered to participate and provides a description of how drill down is used within the OLM, before reporting on the impact on self-assessment accuracy and tasks relating to the regulation of cognition.

8.1 How is an Open Learner Model with Drill Down Approach Used?

Analysis of this sub-question focuses on consistency of use and user acceptance.

8.1.1 Is an open learner model with a drill down approach accepted by its users?

The level of engagement with the OLM technology (Table 29 – Part (i) and Part (ii)) shows that learners answered an average of 223.7 questions each, and the open learner model therefore had a good amount of data to display. <u>The OLM was accessed regularly and on multiple occasions</u> (average 74.93) and each student <u>used the system on 3-16 occasions</u> (mean 6.44). All participants made some use of the drill down functionality, although this was greater for some than for others, as is reflected in the larger standard deviations. The number of questions answered between model inspections was a mean average of 3.51 (SD 1.78) and median of 1.60 (SD 0.86). <u>This shows that there was a bias towards answering only 1 or 2 questions at a time, and also a level of variance in this</u> that is consistent with the proportions of actions shown in the state model in Section 5.6.4 . There are also 320 instances of inspection of the model without answering a question, which may indicate the intervention being used as a reference point.

Table 29: general usage statistics (27 participants).

Usage		Total	Mean	S.D.	Range
(i) Number of Sessions		174	6.44	4.06	3 - 16
(i) Number of questions answered		6040	223.70	213.22	30 - 815
(i) Number of OLM accesses		2023	74.93	75.52	4 - 258
(ii) Inspection of concepts in a topic		2606	95.52	103.18	4 - 326
(ii) Inspection of model for concept		1991	73.74	86.52	0 - 296
(ii) Inspection of beliefs for a concept		440	16.30	17.48	0 - 55
(ii) Inspection of domain for a concept		465	17.2	25.09	0 - 92
(iii) Frequency of inspection, per model inspection	Topic	-	1.23	0.59	0.24 - 3
	Concept	-	0.94	0.71	0 - 3.35
	Beliefs	-	0.22	0.16	0 - 0.73
	Domain	-	0.21	0.16	0 - 0.5

Per model inspection (Table 29 – Part (iii)), the coarser grained elements were accessed more frequently than the finer grained elements (mean frequency 1.23 for topic and 0.94 for concept, versus 0.22 for beliefs and 0.21 for domain). This is significant[14] p>0.0000000001 and is in part a reflection of the structure of the information, as it is required to access concepts before accessing a belief. This does however indicate that not all drill downs were to the deepest level of inspection. The usage level is also far greater than for Study 1 and this analysis confirms that the overall distribution of access of the different features is consistent with Study 1 (in particular the MA(U) group which is the nearest comparison).

[14] Significance calculated with a 1 way ANOVA (related).

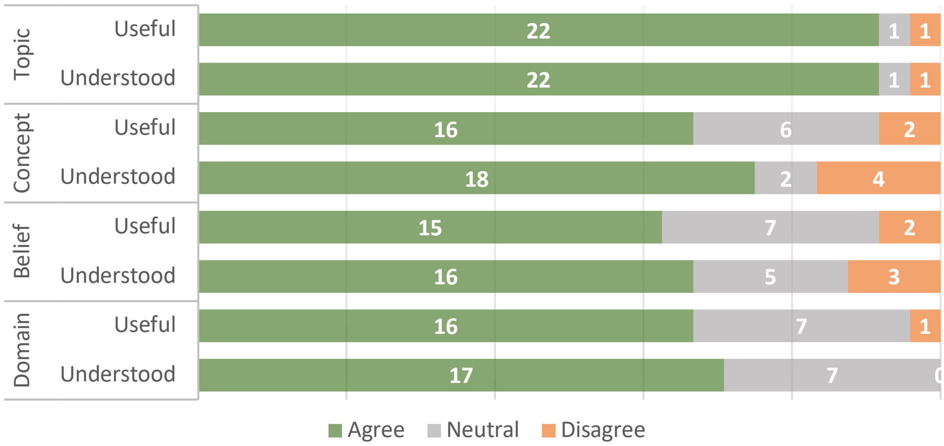


Figure 18: survey responses – intervention is understood/useful (24 respondents).

The log data captured shows regular use of drill down to inspect the OLM in further detail on a conditional basis, and student responses to user acceptance questions on the post usage survey (see Appendix 3) indicated that students understood the learner model facilities available to them and that they generally found them useful (Figure 18). Information at the coarsest level of drill down was indicated to be more readily understood and useful than more detailed elements. The simple overview and breakdown by concept were widely understood and equally useful (1 disagree). When inspecting the concept using a skill meter representation 4 respondents indicated that this was difficult to understand in the context of their interaction, and equally this was the same for inspection of learner model beliefs for 3 participants. The level of user acceptance is slightly greater than as per Study 1, and also confirmed as generally understood and useful, however it still confirmed, as per Study 1 that some students found the finer grained parts of the model of slightly less use and slightly more difficult to interpret, as given by neutral responses in the survey.

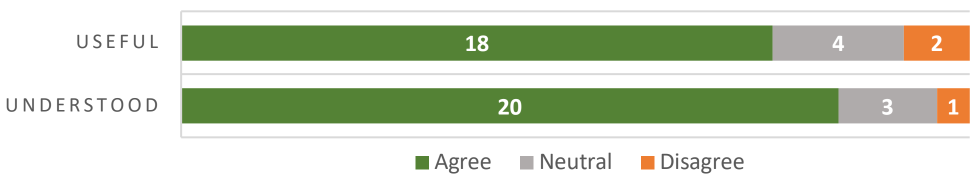


Figure 19: survey responses – drill down approach is understood/useful.

When asked specifically about the use of a drill down structure in the OLM (Figure 19), students generally also found this interpretable and useful (1 disagree, 3 neutral) indicating that the approach overall is generally accepted. The is consistent with Study 1.

8.1.2 Is use of the drill down approach in the open learner model consistent across time?

Interaction across the 8 weeks period incorporated a variety of different interaction approaches with differing levels of usage. A normalised aggregate of the cumulative distributions of these are shown in Figure 20, which indicates that overall within the dataset there is a slightly greater use of detailed inspection at concept level and belief level earlier on in interaction (orange and red line are above the others in the first two quarters).

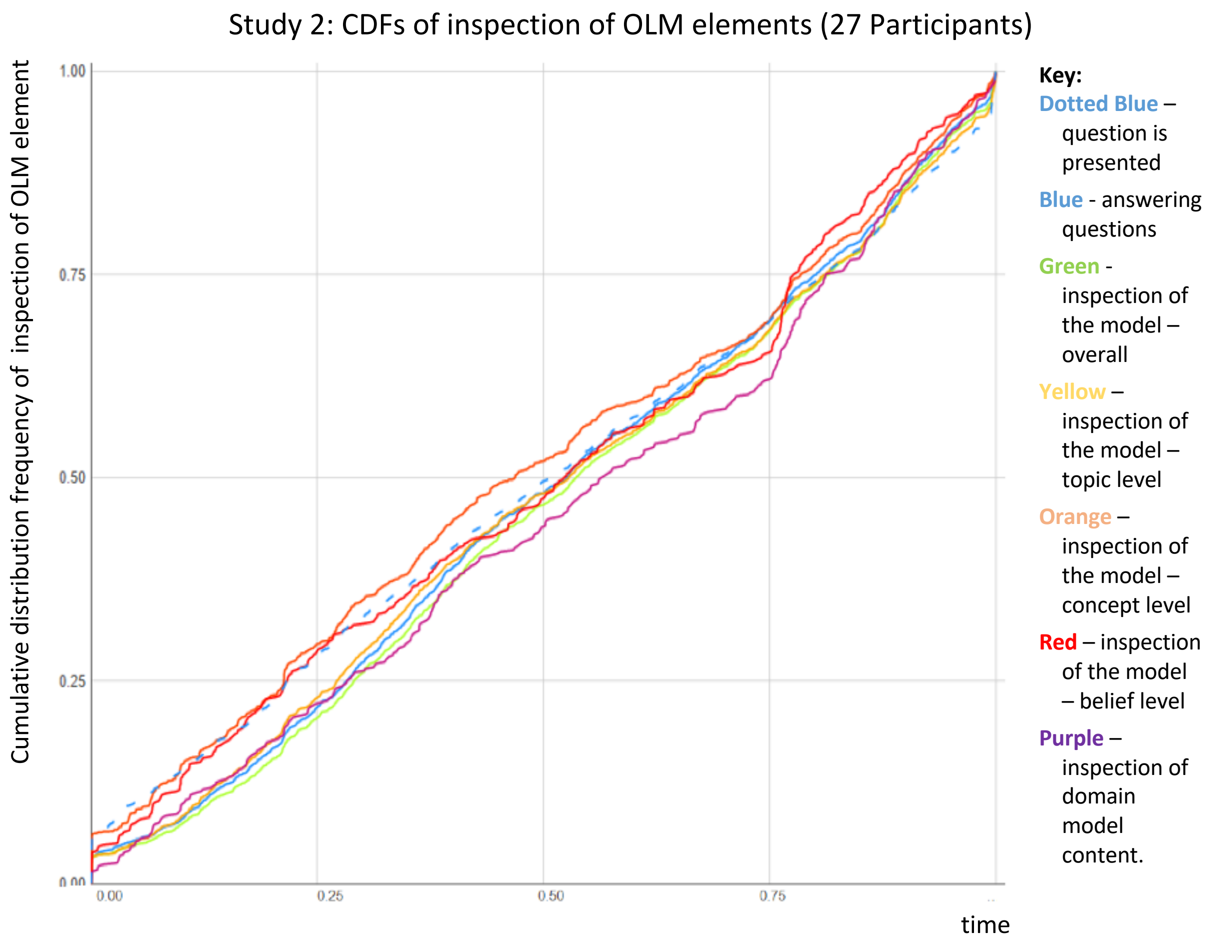


Figure 20: cumulative distribution frequencies of interface usage across time.

The number of unanswered questions is greater in the first half of interaction than in the second half (the blue dotted line is above the blue solid line – <u>this means that students are checking the OLM in the middle of answering questions</u>). The level of interaction and model inspection is quite consistent across the period of the study. As per the analysis in Section 7.1.1 the green line is access of the OLM, and it is the context around which the other lines should be interpreted. It is apparent that topic level inspection (yellow) closely follows this, <u>suggesting consistent and regular use of this across time</u>. In the initial period the concept level (orange) and belief level (red) are used with similar frequency. Access of corresponding domain model information (purple) is latent to this, <u>suggesting a slightly greater use of domain content in latter periods of interaction</u>.

Individual participants' usage patterns across time reveal three broad categories of distribution of equal size (Figure 21 to Figure 23) when considering the access of the more detailed aspects of the model (orange – concept inspection; red – belief inspection; purple – domain inspection). In each case, the items of greatest variance are the deepest levels to which the OLM can be drilled (i.e. learner beliefs, and corresponding domain content). <u>In each group, use of the coarser grained elements of drill down stays fairly consistent</u> with the average access of the model (yellow and orange lines follow the path of the green). It is the <u>deepest level of inspection that have the greatest variance</u> (red line). The access of domain material is usually latent to this (purple line). The cumulative distribution of usage, as presented here, is very similar to the MA(U) group in Study 1. Both have the characteristic of a <u>very linear progression and overall consistent usage across time</u>, particularly with reference to the coarser grained elements of the technology.

Example CDFs of OLM inspection: early use of detailed elements

9 participants fall into the category where the beliefs (red), domain (purple) and to a lesser extent concept (orange) are each notably above the average level of access (green). The more detailed aspects of drill down are used to a greater extent earlier in interaction.

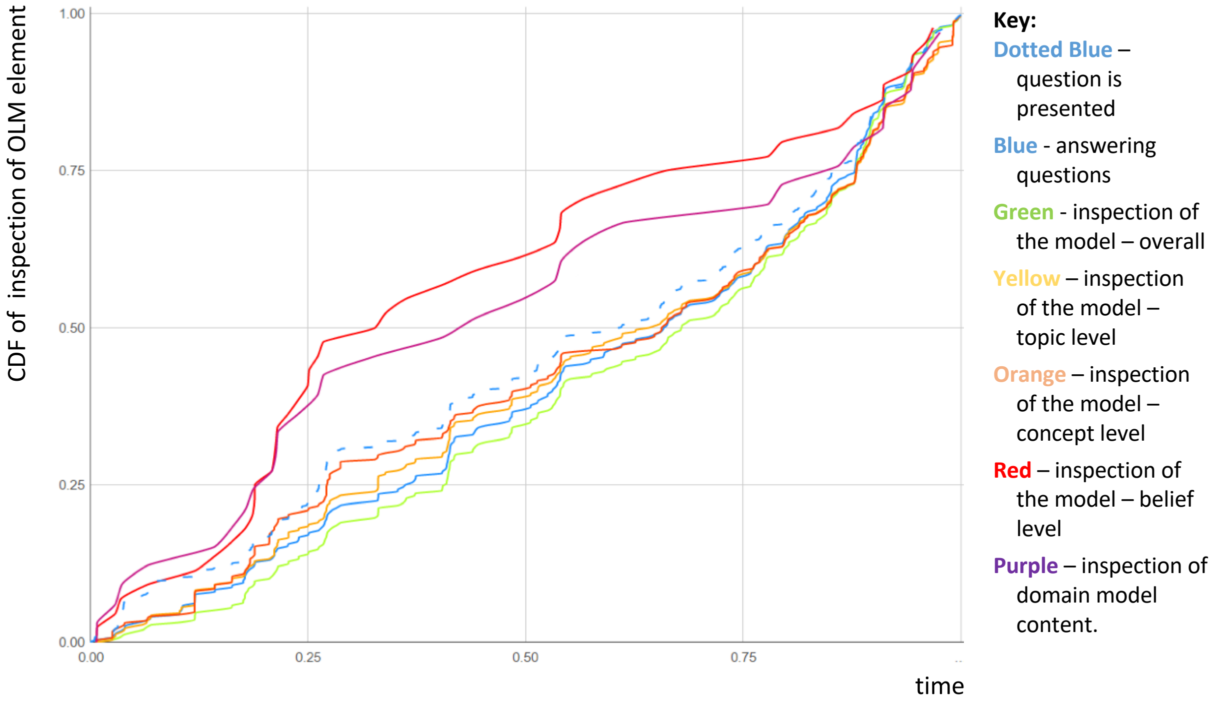


Figure 21: Example CDFs of early use of detailed interface elements.

Example CDFs of OLM inspection: consistent use across time

9 participants fall into the category of fairly consistent usage across time. That is to say the frequency of usage of the more detailed aspects of the drill down capability has similar distribution as compared to general access of the OLM. The differences in the actual numbers of log events generate noise and some deviations from this.

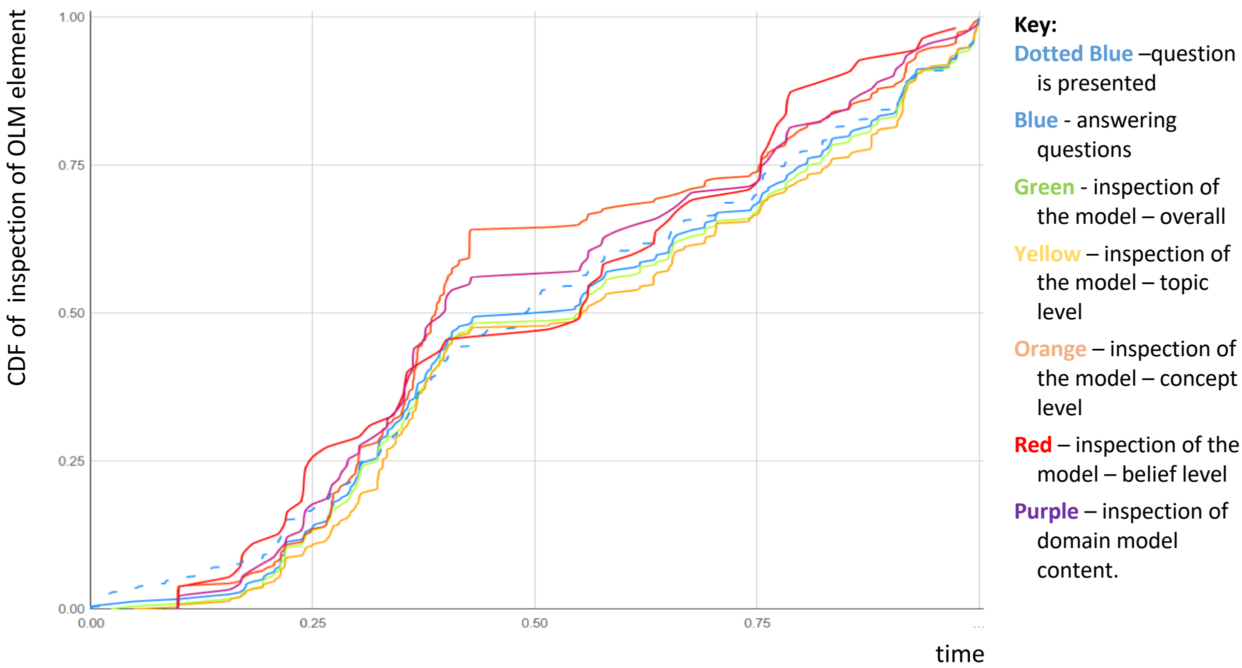


Figure 22: Example CDFs of consistent use across time.

Example CDFs of OLM inspection: later use of detailed elements

9 participants were able to be categorised as using the more detailed aspects of drill down slightly later during their interaction, and thus when there is more information in the model. Notably, concept inspection (orange) is slightly below the OLM assess (green) line, and belief inspection (red) below this. As with the examples of early use of components, inspection of the domain (purple) is latent to inspection of beliefs (red).

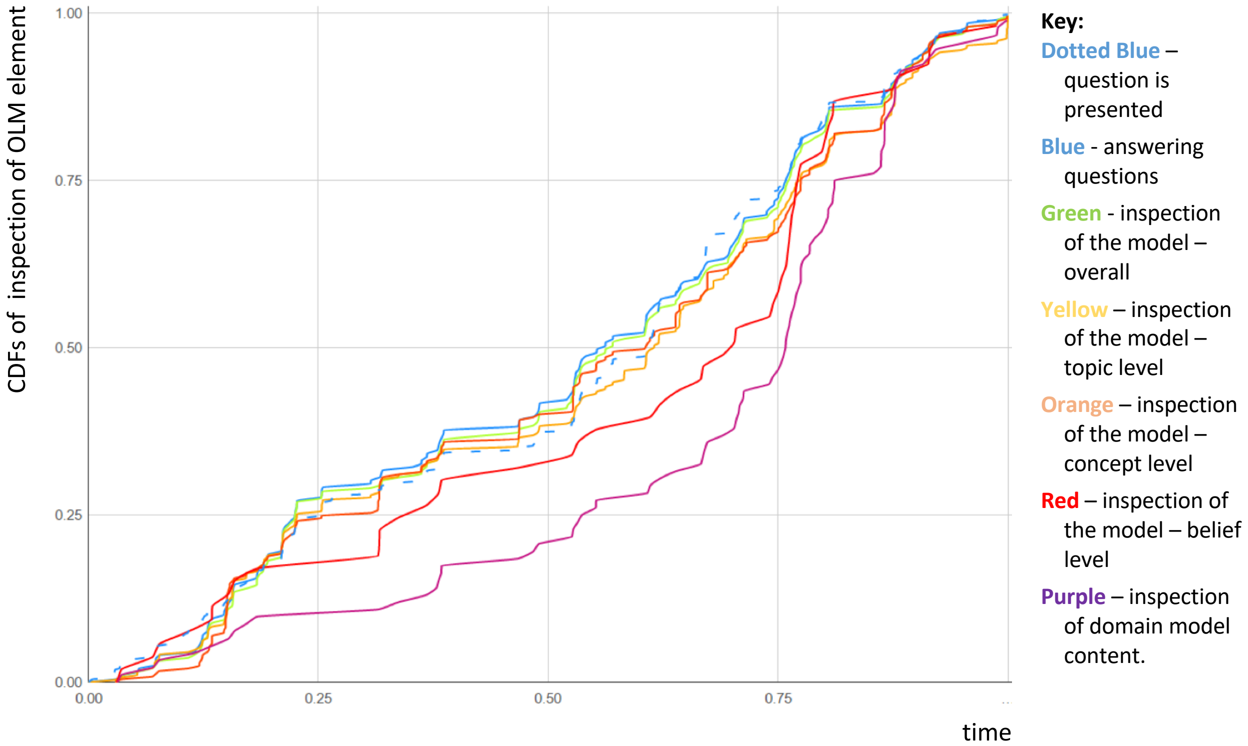


Figure 23: Example CDFs of later use of detailed elements.

Overall, for access of the main coarse grained OLM components this section shows a consistency across time. In terms of whether students find the finer grained elements more useful earlier or later in interaction there is enough variance to say that there may be some individual contextual factors needed to be able to analyse this between participants. Analysis underlines that in both studies, finer grained elements were accessed earlier and later in interaction, even if the frequencies of these do vary.

8.1.3 Is drill down always used when inspecting the open learner model?

Both of the above measures give information on the frequency of use of components at each level of granularity. Table 30 presents the deepest level to which the OLM is drilled on each inspection episode and shows that the majority of sequences go beyond initial surface level of information to more detailed elements of the model. Although there is a lot of variation within the dataset, potentially highlighting students' individual strategies and different levels of usage, the distribution of the probabilities of inspection to each of the levels of detail is consistent between participants (significant[15], $p<0.05$). This is consistent with the MA(U) group, and the general finding in Study 1.

Table 30: frequencies of deepest level levels of OLM inspection.

	Summary	Topic	Concept	Beliefs[16]	Domain
Mean number of sequence occurrences	7.56 (SD 10.47)	21.22 (SD 23.46)	27.37 (SD 43.77)	14.30 (SD 16.63)	15.15 (SD 21.77)

The underlying design of the system is built on the interaction state model of Section 5.6.4 and at each point the learner has a finite number of navigational possibilities, including switching to another focus of the same granularity. Figure 24 shows that most OLM inspections involve the viewing of at least one topic in greater detail (p=0.9) (use of drill down). In approximately half of these cases participants went on to view individual concepts within the topic in more detail – this is the most common path of interaction before deciding to provide more information to update the model, via questioning. In about p=0.3 of cases, participants went on to view in detail either specific beliefs in their learner model (p=0.19) or domain content for the concept (p=0.1) – equally the most detailed level of information available in the OLM. It is interesting to note that in p=0.57 of the cases, after viewing their own beliefs students then inspected the equivalents for the domain; this happens in only p=0.15 of cases in reverse. This shows that students overall had a preference for viewing their own beliefs before the domain model and that both are often used together.

[15] $p<0.05$ - significance calculated with a 1 way ANOVA (related).
[16] Beliefs and domain are to the same level of granularity; sequences will be counted twice if the participant has accessed both of these resources.

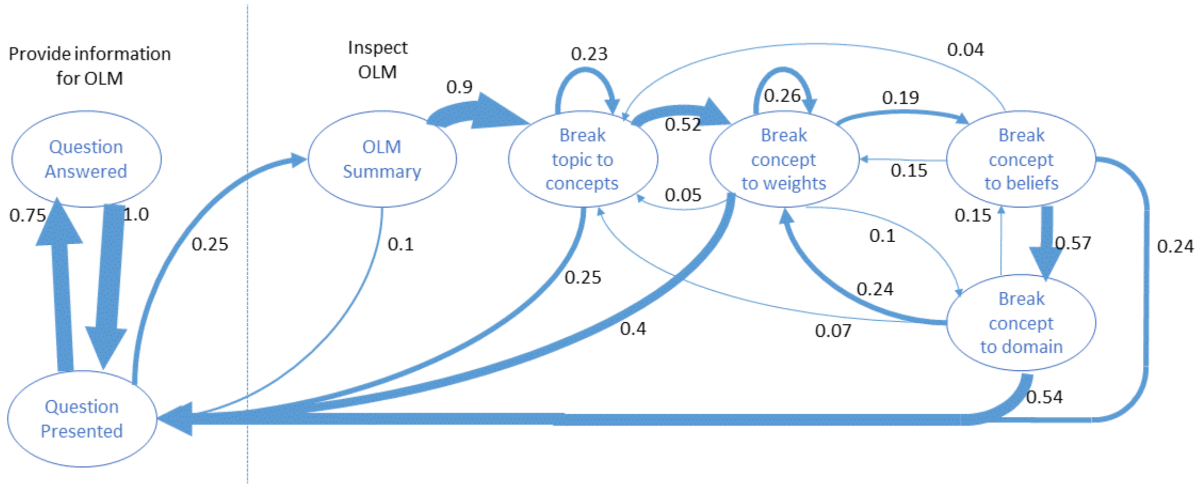


Figure 24: state model for system interaction with the average probabilities of transition.

The overall picture of usage is similar to that which is reported in Study 1 and of greatest similarity to the MA(U) and BO(U) groups. <u>The most common course of interaction is to drill down into the learner model as far as concepts</u>, occasionally going deeper to view beliefs. It also confirms that there is some <u>overview behaviour in terms of skipping through concepts in quick succession</u>, and also that learner beliefs are inspected before domain content. The probability of accessing the OLM is approximately p=0.25, as compared to answering a further question, which brings <u>consistency with Study 1 in terms of the likelihood of making use of the OLM in the context of updating it</u>.

8.1.4 Summary

The analysis presented here shows consistency with Study 1 both for user acceptance of the OLM and the use of a drill down structure in its interface. Student use was consistent across time but with a level of variance at the deepest levels of drill down. This implies some variance in terms of student strategy or conditional use. There is a strong indication that drill down is almost always used in some form when inspecting the model, and that there is evidence to suggest that students are overviewing the model's contents as well as drilling deeper into it on a conditional basis. Consistent with Study 1, students also mostly answered questions in very short bursts, with a bias towards 1-2 questions at a time. Key findings are summarised in Table 31.

Table 31: how is an OLM with a drill down approach used? - key observations.

1(a) Is an open learner model with a drill down approach accepted by its users?	**Yes.** Users indicated their acceptance of each level of drill down and of the drill down structure of the OLM, indicating that both are understood and useful. Students rated the information at the coarsest levels of granularity the most useful. Participants within the sample made good use of the technology on multiple occasions across the period of study.
1(b) Is use of a drill down approach in the open learner model consistent across time?	**Yes.** Drill down use is consistent across the study showing conditional access of finer grained elements of the OLM. There is greatest variance for the deepest levels of drill down, with equal groups of early and late use of the deepest level.
1(c) Is drill down always used when inspecting the open learner model?	**No.** Whilst the majority of accesses of the OLM had a drill down of some sort, the depth and extent of inspection varied, with some indication of conditional aspects of interaction.

8.2 What is the Impact of an Open Learner Model with a Drill Down Approach on Self-Assessment Accuracy?

This thesis investigates the impact an OLM with a drill down approach has on self-assessment through consideration of changes in self-assessment accuracy across interaction. This is indicated by the alignment of student diagnostic and behavioural models and the extent to which this varies when making more intensive use of the drill down functionality in the OLM. The identification of suitable points of measurement for this comparison is detailed in Appendix 5.

8.2.1 Will student self-assessment accuracy increase over time when using an open learner model with a drill down approach?

Table 32 shows that on average there is little change in the accuracy of self-assessment throughout the period of interaction and that there is great variance in the data between participants. <u>The average discrepancy between student self-assessment (diagnostic model)</u>

and the state of the behavioural model is reduced by 4% (not significant). This finding is in alignment with other studies involving students' use of an OLM as a formative assessment opportunity (e.g. Mitrovic and Martin, 2007; Kerly et al., 2008). As compared to the findings of Study 1, the average reduction is similar and also not significant (particularly as compared to the uninstantiated groups). This might suggest that the extended period of time over which the study has taken place has not been an influencing factor, however it is also important to note that the size of the domain content is greater, as it is required to mirror the structure of the students' course.

Table 32: mean discrepancy between the behavioural and diagnostic models, per quarter.

	End of quarter				Difference between
	1	2	3	4	1st and 4th quarters[17]
Mean model state	0.82 (SD 0.11)	0.79 (SD 0.15)	0.80 (SD 0.16)	0.81 (SD 0.17)	-0.01 (SD 0.17)
Mean confidence model	0.81 (SD 0.19)	0.83 (SD 0.17)	0.85 (SD 0.17)	0.85 (SD 0.17)	0.04 (SD 0.15)
Mean discrepancy (absolute)	0.14 (SD 0.12)	0.12 (SD 0.08)	0.12 (SD 0.10)	0.10 (SD 0.09)	-0.04 (SD 0.12)

As outlined in Appendix 5 it is also worth noting that changes in the state of the models are frequent and changes in student confidence (diagnostic model) are likely to be latent to a corresponding change in the behavioural model, if viewing the OLM is a factor that might effect on the self-assessment content. The overall level of completeness of the models (and so model accuracy) increases throughout the period of interaction (to cover the whole domain to a complete level of accuracy at least 4 questions must be answered for each concept – a systematic approach to providing this information is not reflected) therefore there is a further issue here affecting the precision of the initial assessment. If students have also started working with the concepts they found the most straight-forward this may also lead to an overestimation. The literature of Chapter 2 confirms that variance of this type should be expected in self-directed learning.

[17] Not significant – two tailed t-test.

Within the participant sample it is worthy to note that some participants did improve in self-assessment accuracy during the evaluation. As context to this it is considered whether the proxy measure of students indicating their self-assessment ability can identify those whose change in self-assessment accuracy is more likely to be impacted during interaction. As shown in Figure 25, Study 2 endorses the findings of the two instantiated groups in Study 1 that there is a significant[18] correlation ($p<0.05$) between those who increase their self-assessment accuracy and who self-identify as stronger at self-assessment.

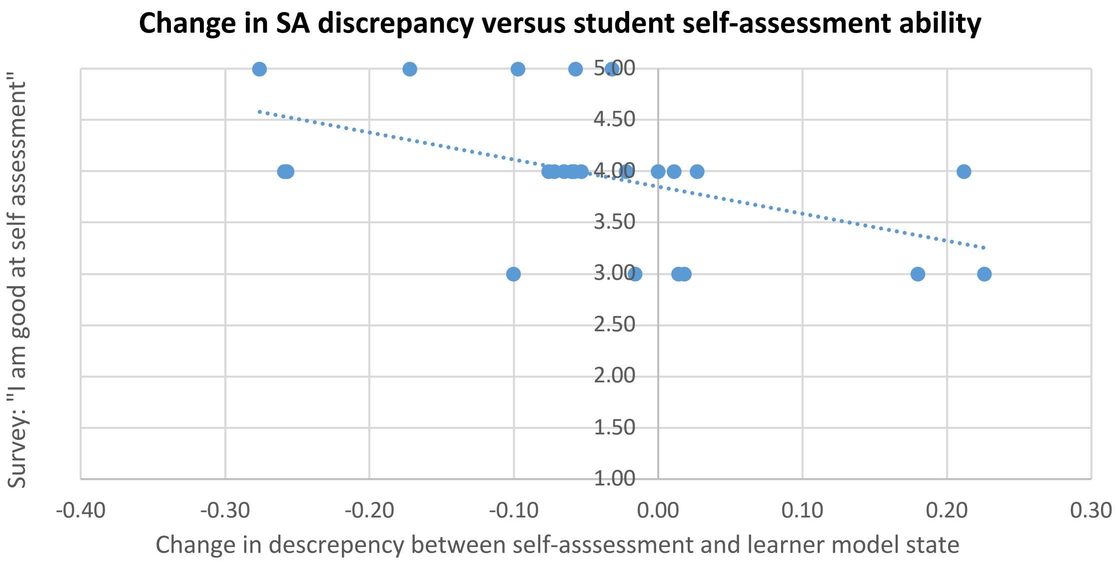


Figure 25: change in self-assessment discrepancy, Vs student self-assessment ability.

Additionally, when considering the context around changes in self-assessment accuracy as compared to the frequency of access of the learner model, Study 2 matches three of the experimental groups from Study 1 in confirming there is no significant[19] correlation at any level of drill down between frequency of OLM access and improvements in self-assessment (Table 33). This is in contrast to the MA(U) group of Study 1 where there was significant correlation. The supporting analysis for this is included in Appendix 7, Section 1.

[18] Significant. $p<0.05$. Spearman rank correlation
[19] Spearman rank correlation.

Table 33: changes in self-assessment discrepancy, compared to frequency of OLM access.

Level of drill down	Correlation between frequency of OLM access and student ability to increase the accuracy of their self-assessment during interaction.		
	Description	R	Significance
Summary	More frequent OLM use, greater increases in accuracy	0.0655	Not significant
Topic	More frequent OLM use, greater increases in accuracy	0.1720	Not significant
Concept	More frequent OLM use, greater increases in accuracy	0.3240	Not significant
Beliefs	More frequent OLM use, greater increases in accuracy	0.1596	Not significant
Domain	More frequent OLM use, greater increases in accuracy	0.2611	Not significant

8.2.2 Will student self-assessment accuracy increase to a greater extent when greater use is made of drill down?

With reference to the deepest level of inspection on each access of the OLM (as per the classification of inspection episodes in Section 5.6.4) there is significant evidence to suggest that <u>students who more frequently drilled down further into the model (to beliefs, domain, or concepts levels) were those who managed to achieve a greater increase in self-assessment accuracy:</u> concept ($p<0.05$); belief ($p<0.05$). (Table 34, with supporting evidence in Appendix 7, Section 2).

Table 34: changes in self-assessment discrepancy, compared to depth of inspection.

% of OLM inspections where the deepest level of drill down was	Correlated with change in discrepancy between student behavioural and diagnostic models over the period of interaction. (Spearman rank correlation)		
	Description	R	Significance
Summary	Decreased accuracy when drilled to summary level	0.1908	Not significant
Topic	Decreased accuracy when drilled top topic level	0.2539	Not significant
Concept	Increased accuracy when drilled to concept level	-0.4316	**p<0.05 (significant)**
Beliefs	Increased accuracy when drilled to belief level	-0.1326	**p<0.05 (significant)**
Domain	Increased accuracy when drilled to domain level	-0.0663	Not significant

The correlations in the data for those who more regularly terminated an episode of learner model inspection (i.e. after viewing only the overall summary or a topic level inspection), might suggest the opposite – that more regular use of a <u>shallower level of inspection leads to a decrease in self-assessment accuracy (not significant)</u>. Study 1 did not present any significant findings for this aspect of the analysis.

8.2.3 Summary

The analysis of this section confirms the finding of Study 1 that there is no generalisable increase in self-assessment accuracy during interaction for all participants, although some participants do increase in accuracy and overall there is a non-significant increase in accuracy. The findings of Study 1 are also confirmed that students indicating their self-assessment ability can be used as a proxy measure to identify those who are more likely to increase in self-assessment accuracy while using an OLM with a drill down approach.

With reference to frequency of use of the OLM, there is a contrary finding to Study 1 that suggests there is no correlation between the frequency of access the OLM and changes in self-assessment accuracy, however there is significant evidence here to confirm the finding of Study 1 that more frequent use of the deeper levels of drill down correlates with larger increases in self-assessment accuracy. Key observations are summarised in Table 35.

Table 35: evidence of metacognitive affect - key observations.

2(a) Will student self-assessment accuracy increase over time when using an open learner model with a drill down approach?	**No.** Consistent with Study 1, there is a non-significant increase in self-assessment accuracy and so this does not generalise amongst participants. As in Study 1, there is significant evidence to suggest that students indicating their self-assessment ability may be used a proxy measure to identify those who will improve self-assessment accuracy to a greater extent.
2(b) Will student self-assessment accuracy increase to a greater extent when greater use is made of drill down?	**Yes.** In contrast to Study 1, there is significant evidence that those who make more frequent use of deeper levels of drill down increase in self-assessment accuracy to a greater extent. This is not related to frequency of access of the OLM.

8.3 What is the impact of an open learner model with a drill down approach on the support for the regulation of cognition?

The thesis now considers participants’ use of the OLM technology in terms of interaction behaviour that might support tasks that are important to the regulation of cognition. This section focuses on students’ inspections of their problems and use of the technology in a focused way as might indicate problem solving, debugging or comprehension behaviour.

8.3.1 Will students use drill down in the OLM to inspect information about problems?

Section 5.6.7.1 describes the calculation for determining the difference between the behavioural model state for an item selected for drill down and the average state of the behavioural model at the point at which the navigation occurred. Figure 26 shows that items inspected in greater detail are for the most part weaker than the average state of the behavioural model at the specific point of access (significant[20] for concepts $p<0.001$, beliefs $p<0.0001$ and domain $p<0.0001$ inspections). As a generalisation, with concepts where the model state is significantly below the mean, students will drill down further to inspect their own beliefs, and with the most problematic of concepts also the domain model. The finding that the deeper the level of drill down the more problematic a concept is (as compared to the behavioural model) is also significantly[21] consistent across participants ($p<0.0000001$) and consistent across the full period of the study. Further supporting detail of the analysis is included in Appendix 7, Section 3.

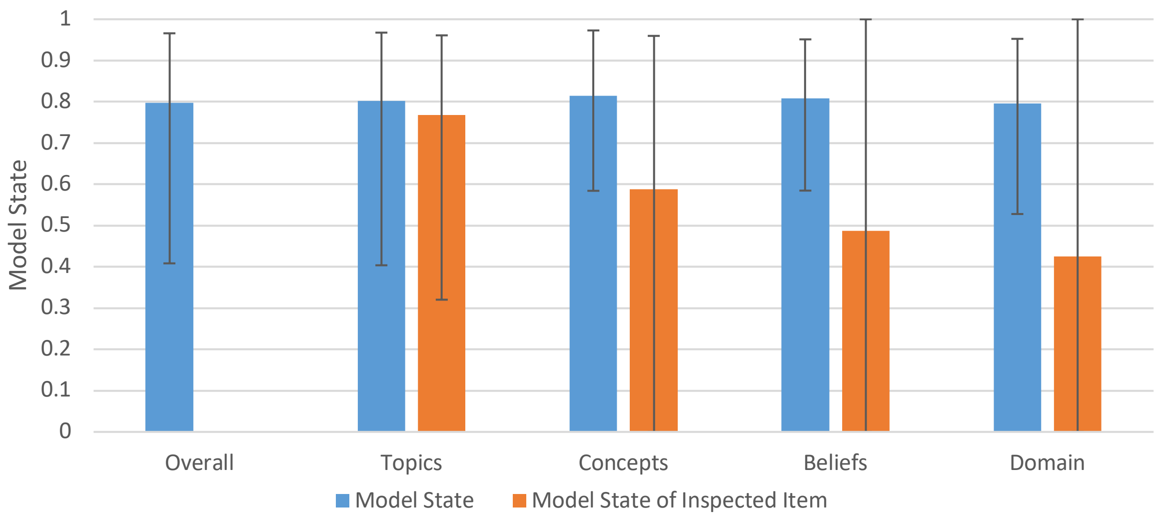


Figure 26: behavioural model state and deviation from the model state at point of access.

[20] Significance between model state and inspection state calculated with a two-tailed T-test.
[21] Significance between drill down levels calculated with a 1 way ANOVA (related).

In the post usage survey (Figure 27), participants were asked for their perceptions of reasons for use, relevant to actions important to the regulation of cognition and reasons for which the model is initially opened to the learner. (See Section 5.5.2 for commentary and Appendix 3 for the survey.)

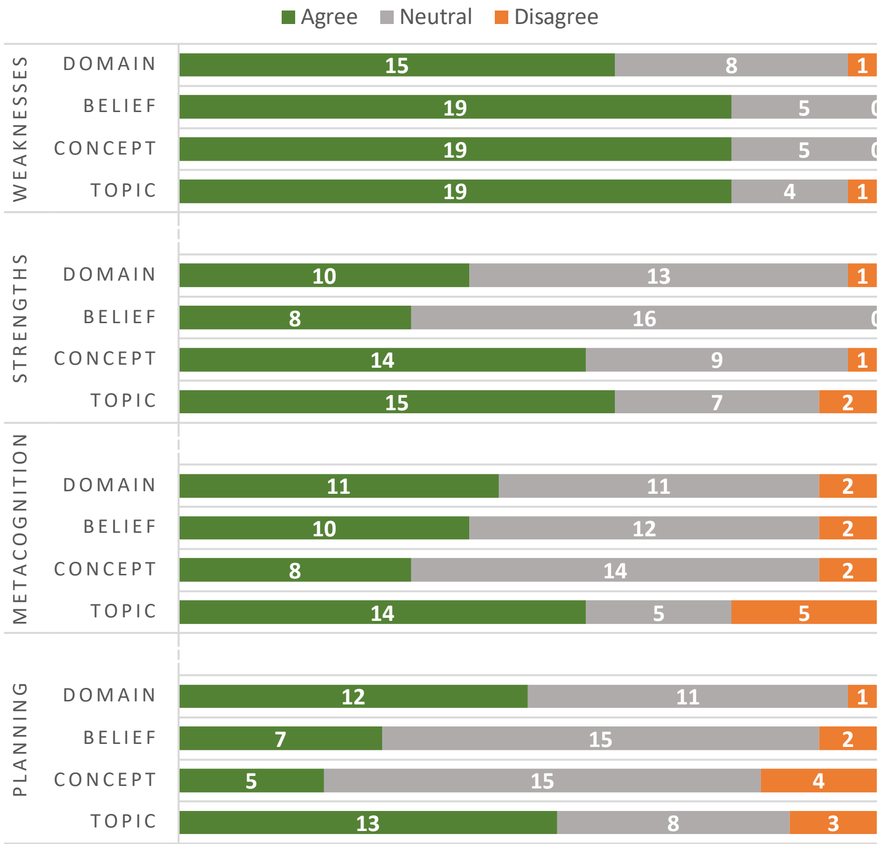


Figure 27: survey items – reasons for intervention use.

The majority of students indicated that they <u>used the OLM for identifying weaknesses, at each level of granularity and this finding is consistent with Study 1</u>. However, in contrast to Study 1 only approximately half of students agreed that they used the OLM for obtaining information about their strengths, and more students considered the coarser grained representations of their model as better for this.

Relevant to actioning the information viewed at each level, students were asked whether they used the information in planning subsequent action. Approximately half of participants agreed that they used the coarser levels of drill down for this purpose, but the majority of participants remained neutral or disagreed that they used the deeper levels of drill down for this. <u>This is in contrast to Study 1</u> where approximately half of participants agreed that the used the OLM for planning at each level of drill down. Students were also asked whether they used the OLM to help them think about how they were approaching their learning of the subject matter (labelled as 'metacognition' in Figure 27). <u>Consistent with Study 1</u>, whilst just under half of participants agreed that they had done this, the coarsest level of drill down was indicated as most use for this purpose.

8.3.2 Will students use drill down in the OLM to inspect information about areas of uncertainty?

Section 5.6.7.2 describes the calculation for determining the difference between the diagnostic model state for an item selected for drill down in the OLM and the average state of the diagnostic model, at the point at which the navigation occurred. Figure 28 shows that <u>while students on average inspected items on which they were slightly less confident, this is not significant[22]. This is also consistent across the different depths of drill down</u>. The supporting evidence is included in Appendix 7, Section 4, and shows that there is a good level of variance between participants in terms of the values of the self-assessment model at the point of inspection. <u>The findings here are consistent with those of Study 1</u> which also showed no significant correlation between use of the OLM and the state of the diagnostic (self-assessment) model.

[22] Significance between model state and inspection state calculated with a two-tailed T-test.

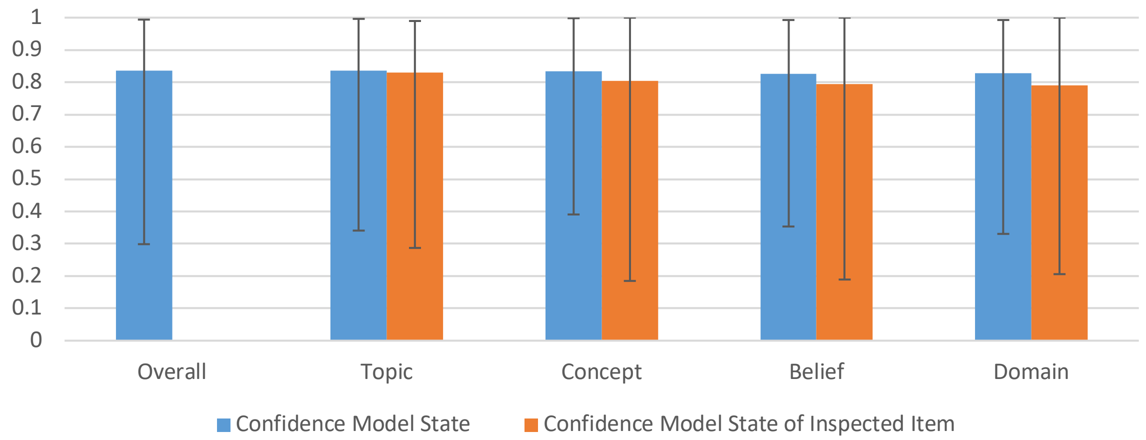
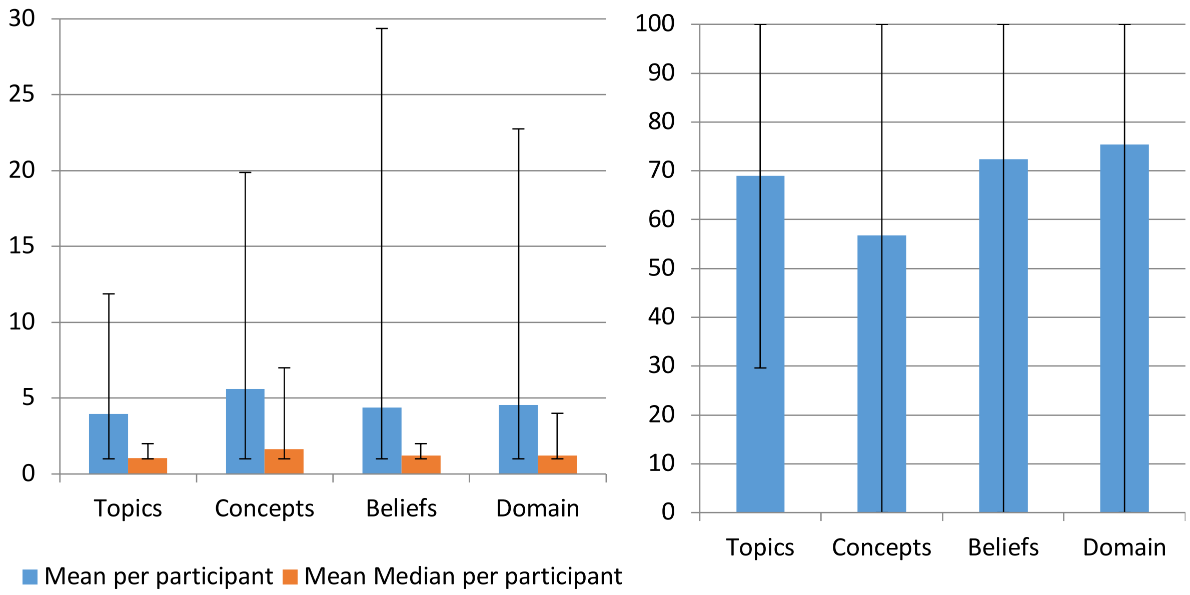


Figure 28: diagnostic model state and deviation from the model state at point of access.

8.3.3 Will students use the OLM with a drill down approach to focus on one domain area at a time?

As per the calculation outlined in Section 5.6.7.3 , Figure 29 shows that across the different levels of drill down the mean values for how recently an item selected for inspection was updated are approximately in the 4-6 updates ago range. However, the distribution of the data shows that there is a strong bias towards the topic or concept being encountered within the last block of questioning (average medians range 1.04 – 1.64). At each level of drill down the majority of topic and concept inspections pertain to the last set of updates (means range 56.8% - 75.3%). This is consistent with the findings of Study 1 and confirms that students are mainly inspecting the last domain area updated but are also revisiting items previously encountered. The non-significant[23] variance across the different levels of drill down suggests that the decision to drill into the model is not affected by how recently the item was updated. This is consistent with three of the experimental groups of Study 1 but contrasts the MA(U) where the deeper the model was drilled, the more recently it had been updated.

[23] Significance between drill down levels calculated with a 1 way ANOVA (related).



(i) mean and median*

(ii) % from last set of questions*

Figure 29: for parts of the model inspected in more detail, how many question-answering sessions ago were the same topic/concept encountered?
*Not significant: 1 way ANOVA (related).

When combined with the state model information of Section 5.6.4 , this potentially shows that that students are using the concept level of drill down to gain an overview and are cycling through related concepts as a reference point, whereas beliefs might be more readily inspected to gain detailed feedback on content that is currently in focus. With regard to the distribution of the information, it might be considered that there are two actions being completed here: searching for feedback on the work currently in focus; and browsing to gain a general overview or establish the state of the model in other domain areas, for other purposes. Additional analysis presented in Appendix 7, Section 5 also confirms that there is significant[24] evidence to suggest that those who increased in self-assessment accuracy to a greater extent were those who also made more regular inspection of items that had not been recently updated.

[24] Spearman rank correlation.

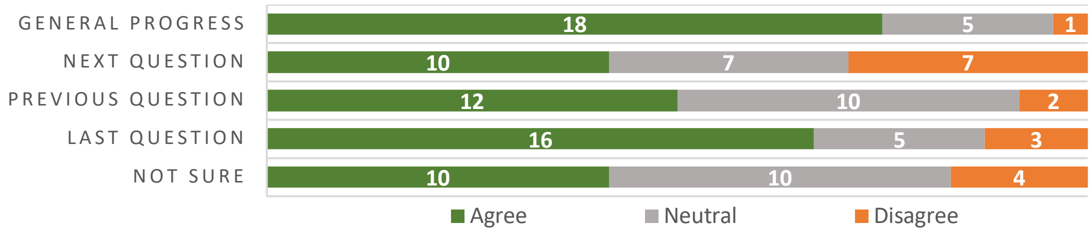


Figure 30: stated reasons for viewing the learner model, in the context of updating it.

In the post usage survey (Appendix 3 and Figure 30) the majority of participants agreed that they inspected the OLM to see general progress and/or to view information related to the last question answered. This is consistent with Study 1. In contrast to three of the experimental groups in Study 1 approximately half agreed that they weren't completely sure as to the purpose of inspecting the OLM at that particular point in time, which indicates that some students didn't always have a conscious purpose to their inspection. With reference to intentions to inspect the OLM in relation to content encountered in the questions, over half agreed that they wished to inspect information about the last question (16 agree, 5 neutral, 3 disagree) and to a lesser extent a previous question (12 agree, 10 neutral, 2 disagree). Students also agreed to a lesser extent that they were using the OLM for a question they were about to work on (10 agree, 7 neutral, 7 disagree), although this feedforward approach was indicated as being completed by some. This indicates that students' intentions were more towards finding out information about the last domain area encountered than the next or previous question, but there were a variety of intentions. This is consistent with Study 1.

8.3.4 Summary

Consistent with the findings of Study 1, there is a strong bias towards students selecting items that are more problematic than the average model state when drilling into the model. This study has additionally highlighted that it is significant that students will drill deeper into the model when problems are greater. The study has also confirmed that students' decisions to drill into the model are not related to their self-assessments of their ability to correctly apply domain knowledge. Their use of the OLM is for detailed inspection of apparent problems and it was confirmed in the post usage survey as a greater reason for inspecting the model than for identifying strengths, planning or thinking about how they're approaching their learning. In alignment with Study 1, the results presented also confirm that there is a strong bias towards students viewing content that has just been updated, and this is as part of a focussed cycle of answering 1-2 questions and then viewing the corresponding area that is updated in the OLM. This baseline of activity is also observed as consistent throughout this study. This pattern of activity is combined with students using the OLM to observe general process, as is confirmed in the post-usage survey and evidenced by students cycling through concepts (also see Section 8.1). The analysis also suggests that the depth to which the model is drilled is not related to how recently it was updated but is conditional on whether it is showing as having identified problems. Key observations are summarised in Table 36.

Table 36: intervention usage compared to the behavioural model - key observations.

3(a) Will students use drill down in the OLM to inspect information about their problems?	**Yes.** Concepts selected to drill down into are generally weaker than the average model state of the student behavioural model. This is significant in the case of concepts ($p<0.001$), beliefs ($p<0.001$) and domain ($p<0.001$) content. The deeper the level of drill down, the more likely it is that students' inspection will be of weaker knowledge level, as compared to the average state of the model (significant, $p<0.0000001$). Students confirmed that inspection/identification of problems is main reason for viewing the OLM. Consistent with Study 1.
3(b) Will students use drill down in the OLM to inspect information about areas of uncertainty?	**No.** There is no significant pattern in the data to suggest that learners' confidence in their abilities relates to their decision to drill into the OLM. Consistent with Study 1.
3(c) Will students use an OLM with a drill down approach to focus on one domain area at a time?	**Yes.** There is a bias towards domain content being selected for drill down that relates to the last update of the model (reported in this section), and that the model is updated in short sharp bursts of 1-2 questions at a time (reported in Section 8.1). The post usage survey confirms that the majority of the participants were consciously desiring feedback on a question just answered and had an interest in general progress.

8.4 Discussion

The research question breaks down into three sub-questions regarding how an OLM with a drill down approach is used, what its impact is on self-assessment and what its impact is on aspects of metacognition. These are now considered in turn.

8.4.1 How is an open learner model with a drill down approach used?

The investigation of Study 2 has reported on a behavioural description of use alongside an 8 week course that has suggested a good level of consistency with Study 1. It has indicated that the OLM, each level of granularity and the drill down approach are accepted by participants who chose to engage with the study. The participants can also be considered as working in a self-directed manner as in using the OLM they have gone beyond the requirements of their course, as suggested by Gabrielle (2016). As in Study 1 variance in student interaction has occurred that would indicate some elements of individual strategies and influence by conditional factors that are each consistent with the literature on individual differences. There is however a common baseline of activity that emerges, which would suggest use for one particular strategy or purpose for the majority of the period of interaction (the consistency of this baseline with the activity of problem solving is discussed further in Section 8.4.3). The presence of a drill down approach in the OLM is intended to support access to a body of detailed information about student understanding when it is conditionally needed by the student. This study has confirmed learners' conditional use of the model is when the OLM shows problems and that this doesn't vary depending on different volumes of new information between inspections of the OLM. This consistency of interaction and the elements of user acceptance considered above allow for an investigation into impact on self-assessment and support for the regulation of cognition with a good level of validity.

8.4.2 What is the impact of an open learner model with a drill down approach on self-assessment accuracy?

Study 2 has confirmed the finding of Study 1 that while some participants did increase in self-assessment accuracy during interaction with an OLM with a drill down approach, there is no generalisable impact on self-assessment accuracy as seen in some other OLM studies (e.g. Mitrovic and Martin 2007; Kerly et al., 2008). Study 2 has also confirmed that students stating their self-assessment ability can be used as a proxy measure to indicate those who are more likely to improve in self-assessment when using this approach. This also correlates with the finding that those who use drill down more often are also more likely to improve in self-assessment accuracy to a greater extent. The OLM may potentially be better supporting stronger learners who already have better developed metacognitive skills, who know when to drill down to get more information from the OLM, and who know how to action OLM information in terms of subsequent self-assessments. It is worthy of note that engagement with the technology inducts students in completing self-assessments, and potentially raises the profile of the action such that it may be applied to future learning episodes or outside the context of the use of the technology – learners were free to complete other activities throughout the 8 weeks. The measurement of change in self-assessment accuracy also relies on the assumption that the OLM, which is being used in a self-directed learning setting, is being used for more than diagnostic assessment (one time period only, per domain area) and that it is systematically returned to at a later point in during the course.

8.4.3 What is the impact of an open learner model with a drill down approach on the support for the regulation of cognition?

Study 2 confirms the findings of Study 1 that students showed particular interest in inspecting areas identified as problematic in greater detail, and that this is influenced by information that the OLM presents, rather than students' confidence in their ability to correctly apply domain knowledge. Building on the narrative of Study 1 (see Section 7.5) this is a conditional factor for drilling down into the model that is consistent with the majority of interaction. When compared to tasks important to the regulation of cognition, students' detailed inspection of problems is a prerequisite step in problem solving and important to tasks such as comprehension and planning. Student use of drill down also emphasises that they are able to complete these actions under their own control without any active navigational element, and this is an important cognitive-based prerequisite for metacognitive activity that is supported by the presence of a drill down structure. A secondary behaviour identified is students completing some level of overviewing of the current state of understanding and drill down appears to play a lesser role in this.

The baseline pattern of interaction identified in both studies is students answering 1-2 questions at a time followed by immediate inspection of the domain area that has just been updated, and conditional use of drill down if it is problematic. Whilst there is some variance in this, which implies a range of tasks and strategies being employed, the consistency of behaviour is an indicator for good practice that will allow students to understand progress, the adaptive nature of the technology, discrepancies and potential misconceptions (Lazarinis and Retalis, 2007), or potential reflection on the state of beliefs rather than a learning trajectory (Van Labeke et al., 2007). The pattern may also indicate behaviour of students attempting to influence the presentation of the model, of identifying why something is not understood, or of attempting to address a problem. This could relate to students attempting to justify (to themselves or the system) or validate their current state of cognition, a pattern which Hartman (1998) states may contribute to metacognitive development. Use of the system in this way could also be interpreted as students attempting to use the OLM for its formative assessment purpose or as a diagnostic test where the feedback from question attempts is being used as 'advice for action' (Whitelock, 2010). The effects of this may also be applied outside of the timeframe of interaction or to different dimension of activity (e.g. whether relating to the person, task, strategy or experience) (Flavell, 1976). There is also potential here that students are also offloading some aspects of metacognition on to the technology, as to not having to complete this themselves. They may be reflecting on what's going on but without this requiring them to build a detailed model of what they're doing – this may be offloaded on to the computer as part of a dialogue.

Most fundamentally the baseline behaviour and its variance, the detailed inspection of problems and the evidence of focussed interaction all align with problem solving strategies such as a trial and error approach to updating the learner model or a hypothesis testing strategy. This in turn implies that the impact of an open learner model with a drill down approach includes support for facilitating tasks important in the regulation of cognition, as part of metacognitive activity.

8.5 Summary

This chapter has reported on a study with 27 learners of basic engineering mathematics at university level over an 8 week period. Results have confirmed the findings of Study 1, have shown user acceptance of the OLM and presented evidence to suggest that students are using it in a focussed and intentional way. There is an inconclusive finding regarding the OLM facilitating improvements in self-assessment accuracy, but a suggestion that it is of benefit to some, and a proxy measure is proposed to identify these students. There is also suggestion that interaction behaviour is consistent with some aspects of tasks important in the regulation of cognition, and in particular students' interest in using drill down with particularly problematic domain areas.

The thesis now moves forward to consider the research contributions, limitations and future work.

Chapter 9

SUMMARY AND FUTURE WORK

This chapter summarises the results of Study 1 and Study 2, further discusses the research question, identifies limitations of the research and suggests future work. It builds on the discussion sections of Chapter 7 and Chapter 8 which have already discussed the results in terms of the literature. It begins with a summary of the main results of the evaluation sections of the thesis and then more formally responds to each of the research sub-questions.

9.1 Summary of the Sub-Research Questions and Observations

This thesis has considered aspects of *how and open learner model with a drill down approach is used and what its impact is on self-assessment and metacognition.* The role of the technology built to investigate this has been to afford learners in the domains of basic engineering mathematics and Boolean logic the opportunity to engage with a body of information about their current state of cognition, in a self-directed learning context. Building on the literature, the research has used a behavioural description of use (evidenced through student interaction) and a post-usage survey (of student perceptions) to consider the research question in three sub-parts:

- How is an open learner model with a drill down approach used?
- What is the impact of an open learner model with a drill down approach on self-assessment accuracy?
- What is the impact of an open learner model with a drill down approach on the support for the regulation of cognition?

Table 37: observations and statistical significances of the results in Chapter 7 to Chapter 8.

Section	Key Results / Commentary	Statistical Significances	Main Implications/Observations
1) How is an open learner model with a drill down approach used?			
a) Is an OLM with a drill down approach accepted by its users?	**Yes.** Both Study 1 and Study 2 confirm user acceptance (used, understood, useful) of each level of drill down and the drill down approach. Study 2 indicates coarser levels of drill down having higher perceived usefulness.	• *More qualitative and state model driven descriptions of use are given (e.g. proportions of interaction to indicate behaviour, Likert scales).* • Study 2: use of drill down and frequency of OLM inspection is consistent across participants **p<0.0000000001.** • Study 2: drill down is most common to a medium level of granularity, with an overview the second most commonly used and fine grained inspection of the OLM third **p<0.05.**	• User acceptance gives a level of validity to measures used to determine the impact of the OLM on self-assessment and regulation of cognition. • Some variances throughout interaction imply a range of strategies are in place, but there is a baseline pattern of activity to suggest use for one purpose for the majority of interaction. • Drill down use is conditional (sub-question 3 considers this in further detail). • Usage suggests small modular or focussed working when using the technology, and that drill down is being used conditionally when more information is required. This is consistent with the purpose of the principle of drill down use.
b) Is use of the drill down approach in the OLM consistent across time?	**Yes.** Both Study 1 and Study 2 show consistent use across time and there is a baseline pattern of behaviour, in addition to some variance indicating individual strategies. Students work with small bursts of questions and inspect the same domain area afterwards. There is some evidence of students overviewing the content of the OLM, and evidence of conditional deep drilling of the model.		
c) Is drill down always used when inspecting the OLM?	**No.** Both Study 1 and Study 2 confirm that while most accesses of the OLM involve a drill down of some sort, the depth and extent to which the OLM is drilled varies and appears to have conditional elements.		
2) What is the impact of an open learner model with a drill down approach on self-assessment accuracy?			
a) Will student self-assessment accuracy increase over time when using an OLM with a drill down approach?	**No.** Both Study 1 and Study 2 report average non-significant increase in self-assessment accuracy during interaction. Some participants improved, whilst others did not. A proxy measure of students indicating their self-assessment ability may be used to identify those who will improve to a greater extent.	• Students who increased self-assessment accuracy to a greater extent identified as stronger at self-assessment: Study 1: **p<0.05, p<0.01, p<0.05, p<0.05.** Study 2: **p<0.05.** • Study 1: More regular use of the OLM correlates with greater increases in self-assessment accuracy MA(U) **p<0.001.** • Study 1: More regular use of drill down correlates with: (i) greater increases in self-assessment accuracy MA(U) **p<0.05;** (ii) greater decreases	• If students use an OLM with a drill down approach in a way that is consistent with the baseline pattern of activity stated in sub-question 1 and sub-question 3, this alone does not appear to impact on a change in student self-assessment accuracy over the experimental period used in this thesis. • Students indicating their self-assessment ability can be used as a measure to identify those who would be able to increase the accuracy of their self-assessments to a greater extent. • Students who use the deeper levels of drill down more frequently are also more likely to increase in self-assessment accuracy to a greater extent. Study 2 and
b) Will student self-assessment accuracy increase to a greater extent when greater use if made of drill down?	**Yes.** Study 2 and the MA(U) group from Study 1 indicate that more frequent use of deeper levels of drill down correlates with greater increases in self-assessment accuracy, potentially also correlating with the profile of a stronger learner. Study 1 also reports that		

	more frequent use of the OLM relates to greater increases in self-assessment accuracy. The BO(U) group from Study 1 indicates a contrary finding for the deepest level of drill down but is considered as of slightly lower validity as this is not directly used alongside a module of study.	in self-assessment accuracy BO(U) **p<0.01, p<0.01.** • Study 2: Those who used the deepest levels of drill down more frequently increased in self-assessment accuracy to a greater extent **p<0.05, p<0.05.**	three of the experimental groups from Study 1 indicate that this is not related to the frequency of use of the OLM. Overall, this could relate to part of a profile of a stronger learner with better pre-existing metacognitive skills.
3) What is the impact of an open learner model with a drill down approach on the support for the regulation of cognition?			
a) Will students use drill down in the OLM to inspect information about problems?	**Yes.** Both Study 1 and Study 2 are consistent and state that the areas of the domain selected for drill down are on average significantly weaker than the average state of the behavioural model, indicating students are using it to mainly inspect problems in detail. Study 2 additionally finds that the deeper the model is drilled the greater the extent of the problems for the domain area.	• Topic level inspections are for information of average model state. Information at the lower levels of drill down is significantly weaker than the average model state Study 2: **p<0.001, p<0.001, p<0.001**; Study 1: MA(U) **p<0.01,** BO(U) **p<0.05,** MA(I) **p<0.01.** • Study 2: students will drill deeper when problems are of greater magnitude **p<0.0000001.**	• Students are completing detailed inspections of the OLM for areas that are identified as problematic. This is consistent with some of the initial steps in problem solving strategies and useful with regard to planning and comprehension.
b) Will students use drill down in the OLM to inspect information about areas of uncertainty?	**No.** Study 1 and Study 2 confirm that there is no significant pattern in the data to suggest learners' confidence in their ability to correctly apply domain knowledge relates to their decision to drill into the OLM.	• Not significant.	• Students' inspections of their problems is distinct from their perceptions of how strong their understanding is. Their navigational decision is more readily influenced by the open representation of the behavioural learner model.
c) Will students use an OLM with a drill down approach to focus on one domain area at a time?	**Yes.** Study 1 and Study 2 both report a baseline interaction pattern of students answering short bursts of questions and then immediately inspecting the corresponding part of the domain in the OLM, with conditional drilling into the model when it shows problems. This is repeated consistently throughout interaction and is evidence towards systematic and focussed working, and students indicate that they are using it for this purpose, as well as to gain a general overview.	• Study 1: the deeper the model is drilled, the more recent was the update of the item MA(U) **p<0.001.**	• Additional to the task of overviewing the model's contents, when students' focussed working and interest in their problems is combined with the baseline activity pattern this shows a level of consistency with strategies such as trial and error, debugging or hypothesis testing, which are problem solving strategies.

Before responding to each of these aspects of the investigation, Table 37 summarises the main observations and results of Chapter 7 and Chapter 8, and highlights items of analysis that have shown statistically significant results.

The main points from the analysis include:

- The variances in the data imply that there are a range of strategies being used, and these include overviewing of concepts and detailed inspection of problems.
- The model is drilled deeper when problems are of greater magnitude.
- There is a baseline of activity that accounts for the majority of student interaction. It comprises answering short numbers of questions (1-2) followed by immediate inspection of the corresponding domain area in the OLM. Drill down is used conditionally.
- There is no universal impact on self-assessment in terms of students increasing the accuracy of self-assessments during interaction, however students indicating their self-assessment abilities is sufficiently accurate to identify learners who will improve to a greater extent.
- Those who improve in self-assessment accuracy to a greater extent make more regular inspection of the deeper levels of drill down.
- Students' decision to drill deeper into the model is more readily influenced by the presentation of the OLM than by their confidence/perceptions of their ability to apply domain knowledge.
- Students' baseline interaction and conditional detailed inspection of problems are consistent with trial and error or hypothesis testing behaviour, which are problem solving techniques that indicate the presence of the regulation of cognition, as part of metacognition.

Overall, there is support for the OLM being accepted by its users and being used consistently across time. There is an emergent baseline activity pattern and evidence suggesting that a specific use of the technology is for inspecting problems in detail. The drill down structure particularly supports conditional access to this information, and students have shown an element of critical thinking in being able to identify problems. This impacts on increases in self-assessment accuracy for some participants, particularly those who make greater use of drill down, but this is not a universal affect. Student activity also impacts on the regulation of cognition by facilitating behaviour consistent with problem solving strategies, which are metacognitive by definition.

9.2 Response to Sub-Questions

The research question breaks down into a consideration of how an OLM with a drill down approach is used (with emphasis on user acceptance, consistency of interaction and generalisable interaction), its impact on self-assessment (in particular self-assessment accuracy, as in previous OLM studies), and its impact on metacognition (in particular support for tasks important in the regulation of cognition aspect of this).

9.2.1 How is an open learner model with a drill down approach used?

It is appropriate to ask how an open learner model with a drill down approach is used as there is a gap in the literature in this area. Learner models are opened to learners for a range of metacognitive and cognitive benefits, and many often implement a drill down structure in their design as a practical solution to structuring information, however the use of the drill down structure itself is seldom evaluated. Building on evaluation approaches that are common in OLM research the thesis aimed to consider how an OLM with a drill down approach is used with reference to user acceptance and consistent or generalisable interaction. This is done with specific reference to how the use of drill down sits with reference to students' primary interaction task, which is the inspection of the externalised form of the learner model. This thesis followed three lines enquiry which are answered as follows:

Q1(a) Is an open learner model with a drill down approach accepted by its users?

Yes, an OLM with a drill down approach is accepted by its users, in the field of electrical, electronic and systems engineering. Those who chose to engage[25] with the technology engaged across the period of the study and generally indicated the OLM and its drill down structure as understood and useful.

Q1(b) Is use of the drill down approach in the open learner model consistent across time?

Yes, an OLM with a drill down approach is used consistently across the period of the study with the emergence of a baseline pattern of activity (working with small bursts of questions followed by inspection of the newly updated domain area in the OLM, with conditional drilling of the model if this is problematic), and also some variance to suggest that other strategies are also being used (e.g. overviewing of information).

Q1(c) Is drill down always used when inspecting the open learner model?

No, drill down is not always used when inspecting the learner model. The majority of interaction suggests that this is conditional based on whether the behavioural model shows problems, and this is not related students' confidence in their ability to correctly apply domain knowledge. Whilst most OLM inspections have some level of drill down, this appears to be highly conditional.

The evidence of these lines of enquiry suggest that an open learner model with a drill down approach is used consistently across student use with a baseline pattern of activity of short bursts of questions followed by immediate inspection of the area of the OLM that is updated. Drill down is used conditionally to allow students to complete detailed inspection of problems that are apparent. The drill down structure also allows students to gain an overview of their learning without requiring all the detail of the model.

[25] See Section 5.4 and Section 5.8 for more information about the placement of the evaluation within the School of Electronic, Electrical and Systems Engineering at the University of Birmingham. User acceptance considered here does not report on those who did not participate in the study. These numbers are not known.

9.2.2 What is the impact of an open learner model with a drill down approach on self-assessment accuracy?

It is appropriate to ask how an OLM with a drill down approach impacts on self-assessment accuracy as: this is a primary aim of opening the learner model; it is something that is an ongoing line of investigation in OLM research; and it has not been investigated with specific reference to a drill down approach. It is valid to continue to identify characteristics of OLM that are optimal for facilitating good self-assessment practice and fostering self-assessment skills. The literature has identified that the impact of self-assessment accuracy is an important aspect of this and previously impacts have been observed with simple OLMs and students of different ages. It is suggested an appropriate way to measure this is through comparison of learner behavioural and diagnostic models across different points in time. The novel component in this research is the presence of a drill down structure, and so this should be correlated with changes in self-assessment accuracy. The thesis followed two lines of enquiry to answer this:

Q2(a) Will student self-assessment accuracy increase over time when using an open learner model with a drill down approach?

No, student self-assessment accuracy does not significantly increase over time when using an OLM with a drill down approach. Whilst some learners do improve (as can be proxied by students giving an indication of their self-assessment ability) this is not a universal effect.

Q2(b) Will student self-assessment accuracy increase to a greater extent when greater use is made of drill down?

Yes, there is some evidence to suggest that students who make greater use of drill down during interaction will increase in self-assessment accuracy to a greater extent. This potentially could be part of the interaction profile of a stronger learner with better pre-existing metacognitive skills.

These two lines of enquiry suggest that the impact of an OLM with a drill down approach on self-assessment accuracy is that while there is no universal increase in self-assessment accuracy, some participants do improve in self-assessment accuracy during use and these students make greater use of drill down. These learners can also be accurately identified by asking students to give an indication of their own self-assessment abilities.

9.2.3 What is the impact of an open learner model with a drill down approach on the support for the regulation of cognition?

It is appropriate to ask what the impact of an OLM with a drill down approach is on support for the regulation of cognition as metacognitive benefit is one of the primary aims of opening the learner model to students. Metacognitive benefit is an ongoing line of enquiry in OLM research and has not been investigated with specific reference to a drill down approach, it is therefore valid to seek new settings in which OLMs have educational benefit. As metacognition is very much an internalised process it is appropriate that this is first investigated with more tractable aspects of metacognition, which Shraw and Dennison (1994) summarise as including the regulation of cognition, including subprocesses that facilitate the control of learning. These include comprehension monitoring, problem solving/debugging strategies and aspects of planning. Mayer (2013) also extends this to indicate, with particular reference to problem solving, that this has measurable components in: students recognising/identifying problems as distinct from their own perceptions of understanding; structuring problems through detailed inspection and information gathering; the execution of strategies (e.g. hypothesis testing or trial and error); and then monitoring of feedback of the action. This thesis in the context of looking for impact on metacognition focuses on support for the regulation of cognition as evidenced through behaviour consistent with problem solving and comprehension monitoring. Combined with the presence of a drill down approach, the thesis has therefore investigated the following lines of enquiry:

Q3(a) Will students use drill down in the OLM to inspect information about problems?

Yes, students drill into the model mainly in domain areas where problems are identified, and they will drill deeper into the model when the problems are greater. There is evidence of students performing detailed inspections of their problems in a way that is consistent with phases of problem solving, including information gathering, the monitoring/comprehension of feedback, and the execution of problem solving strategies.

Q3(b) Will students use drill down in the OLM to inspect information about areas of uncertainty?

No, students' decision to drill into the OLM is not related to their certainty in their ability to correctly apply domain knowledge. Students' decision to drill into the OLM is more readily influenced by the state of the (behavioural) model that is presented to them.

Q3(c) Will students use an OLM with a drill down approach to focus on one domain area at a time?

Yes, students use an OLM with a drill down approach to complete focused interaction, working with one domain area at a time. This is part of the baseline activity of short question bursts followed by immediate inspection of the same domain area in the OLM. Whilst there is evidence of students also completing other tasks, this pattern accounts for the majority of interaction, and is consistent across participants. This is also consistent with trial and error or hypothesis testing strategies as part of problem solving, and also evidences systematic monitoring of feedback.

These three lines of enquiry indicate that the impact of an OLM with a drill down approach on the support for the regulation of cognition is that it can readily facilitate and evidence student actions that are consistent with problem solving, as a subprocess that facilitates the control of learning (metacognition). Problem solving behaviour is part of the baseline interaction of students interacting with an OLM of this type and drill down is used in a way as to support the conditional access of information that is required specifically when the detailed inspection of problems is to be undertaken.

9.3 Limitations to the Findings

In critical evaluation of the results presented here, a key limitation of the studies includes the extent of the data captured. This thesis has presented evaluation with a total of 87 participants, of which 27 used the technology alongside a course of study over an 8 week period. Evaluation with a larger sample of participants, and in further domains, would help verify the findings, and would allow for inferential statistics to be applied to sub-groups. The analysis may also be extended to include analytics displaying further measures of the generalised behaviour – this thesis has considered a subset of behavioural aspects of procedural and conditional learner interaction. There has also been a challenge of managing a controlled experiment in a real learning setting, where self-directed learning is the literature-suggested best practice approach for facilitating metacognition, tasks regulatory to cognition and practices contributing to metacognitive development such as self-assessment; there are likely confounds that have not been fully controlled, particularly in the context of external student actions. The evaluation participants were all of university age and actively enrolled in university courses in the field of electronic, electrical and systems engineering, this also further limits the findings to this demographic. Similarly, a full learner profile is not known about participants, with particular reference to how interaction links to other aspects of their learning, over time, of which cross referencing and contextualising are potentially important in understanding the presence of any generalised affect.

With the longer-term study there were a series of factors which could not be controlled. These centred on students being free to interact with any other learning resources and actors over the period, variance with respect to the extent of engagement with the technology, and whether weekly use was regular enough for measuring issues related to self-assessment and the regulation of cognition. A further limitation of the data is that other measures to capture aspects of learning are not applied to the evaluation, such as through planning documents, setting and attainment of goals, interviews, or think aloud protocols at regular points during interaction. The decision was taken not to use these collection methods as not to bias activity during interaction, as these may also be considered as intervention techniques in learning according to the literature of Chapter 2.

The design of the technology originates from available literature, however only one design of learner model interface is ultimately evaluated. Designs which allow for the representation of only coarse granularity information, or delayed feedback may activate stronger educational affects (e.g. Shute, 2008; Kay and Lum, 2005). The findings here are, as such, limited to an open learner model, with a drill down approach, in a self-directed learning context, at university level, and in an engineering domain.

9.4 Key Contributions

This thesis contains information of interest to different audiences, including researchers of educational technology, of learning analytics and open learner models, of education and pedagogy, and also those that design and implement technology. Contributions to each of these primary and secondary stakeholders are summarised in Figure 31.

SECONDARY

Technology Designers

- Technology specification/ design and implementation for supporting metacognition.
- Approaches such as drill down are important to include with respect to supporting regulation of cognition, particularly when designing OLMs.
- Design of a method for formative assessment in the domain of mathematics, using a template-based approach to question generation.

Educational and Pedagogical

- OLMs, as a technology designed around metacognitive opportunity remain relevant in the research of technology-based formative assessment benefits.

PRIMARY

Students interact with an OLM with a drill down approach in a way consistent with tasks important to the regulation of cognition, as part of metacognition. Whilst this interaction does not always allow students to increase in self-assessment accuracy, those who will improve can be identified by the proxy measure of them indicating their self-assessment ability.

Researchers

- Proposed ways of measuring metacognitive aspects of student behaviour through interaction pattern analysis and the comparison of different model types, within the constraints/confounds of avoiding influencing metacognitive behaviour.

LA/OLM Practitioners

- Definition of a baseline of activity for students' interaction with an OLM, in the context of providing new evidence for it to be updated.
- Measurement methods as a potential way to update a LM of metacognitive influence or participation in the regulation of cognition.
- Drill down interface approach is relevant to OLMs/LA from the perspective of metacognition.
- OLMs permit interaction behaviour consistent with supporting metacognition and can be designed around requirements that allow for the support of metacognition.

Figure 31: key contributions to primary and secondary stakeholders.

The main finding indicates that the use of a drill down approach in an open learner model can appropriately structure information to support students in their conditional inspection of their problems, in addition to identifying general progress. It supports and evidences interaction that is consistent with problem solving behaviour, which is evidence of the regulation of cognition as part of metacognition. The presence of the drill down approach also supports learner interaction for those who would be more likely to improve in self-assessment accuracy, although the technology itself does not allow for this to have an impact on all learners. For those with an educational interest, this also confirms that OLMs remain relevant in educational research. The thesis contributes to the definition of a baseline of activity that that is representative of students' interaction with an OLM, in the context of the action of providing new information to update it. The thesis proposes a series of evaluation and analysis methods to look at aspects of metacognition without directly aiming to affect it, and these methods may be of further interest to learner model and learning analytics researchers, in the enhancement of how learner models may be constructed. This thesis also highlights that a drill down approach in an interface for technology such as OLMs is of relevance to metacognition and further proposes a technology specification for those that implement technology with metacognition and the research of metacognition in mind.

9.5 Future Work

Much that can be suggested as future work is to overcome the limitations of the studies undertaken, as to further verify the findings. As such, the following may be appropriate:

- Application of the evaluation in a different domain, with learners of a different demographic.
- Repetition with a larger number of participants.
- Cross referencing against a more established learner profile, and access to other resources.
- Use of other techniques, such as planning documents or regular interviews, to monitor a greater spectrum of metacognitive aspects of learning.
- Evaluation with other implementations of intervention methods, such as those with different granular specifications, absence of drill down approach, delayed feedback etc.

As a further point future studies could investigate the effectiveness of using different drill down approaches (such as fish-eye type zooming, modal dialogues, treemaps etc.) and complete this in terms of metacognitive support, and support for cognitive tasks as initiation points for metacognition.

Future work could also consider a categorisation of learner interaction patterns with the technology, such as may be cross referenced against the completion of specific tasks that support metacognition (such as identifying the defined presence of critiquing, planning, questioning, justifying, explaining, knowledge acquisition, knowledge application, comprehension, critical thinking or problems solving (see Hartman, 1998; Boud and Molloy, 2013)), and to consider an evaluation of the effectiveness of each of these against the use of a technology of this type.

9.6 Summary

This thesis has considered issues relating to *how and open learner model with a drill down approach is used and what its impact is on self-assessment and metacognition.* Literature has suggested that an open learner model using a drill down approach is appropriate for learners working in a self-directed context, and that support for student self-assessment and metacognition are reasons for opening the learner model to students. The thesis has shown that a drill down approach evidences and supports a baseline interaction behaviour that is consistent with problem solving and tasks important in the regulation of cognition. It has also indicated that an OLM with a drill down approach can support improvements in self-assessment accuracy for some learners, and those who will benefit to a greater extent may be accurately identified by asking students to give an indication of their self-assessment abilities.

LIST OF REFERENCES

Abrami, P.C., Bures, E.M., Idan, E., Meyer, E., Venkatesh, V. & Wade, A., 2013. Electronic portfolio encouraging active and reflective learning. In: R. Azevedo & V. Aleven, eds. 2013. *International Handbook of Metacognition and Learning Technologies*. New York, NY: Springer. pp. 503–515.

Ahmad, N. and Bull, S., 2008. Do Students Trust Their Open Learner Models? In *International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems*. Springer Berlin Heidelberg. pp. 255-258.

Alotaibi, M. and Bull, S., 2012. Using Facebook to support student collaboration and discussion of their open learner models. In proceedings of *Workshop on Web 2.0 Tools, Methodology, and Services for Enhancing Intelligent Tutoring Systems, Intelligent Tutoring Systems 2012*, 14th-18th June 2012, Crete, Greece.

Anderson, N.J., 2008. Metacognition and good language learners. In: C. Griffiths, ed. 2008. *Lessons from Good Language Learners (Cambridge Language Teaching Library).* Cambridge: Cambridge University Press. pp.99-109.

Aronson, S., Mahanta, L., Ros, L.L., Clark, E., Babb, L., Oates, M., Rehm, H. and Lebo, M., 2016. Information technology support for clinical genetic testing within an Academic Medical Center. *Journal of Personalized Medicine 6(1)*, p.4.

Ashton, H.S., Beevers, C.E., Korabinski, A.A. and Youngson, M.A., 2006. Incorporating partial credit in computer-aided assessment of Mathematics in secondary education. *British Journal of Educational Technology 37(1)*, pp.93-119.

Azevedo, R., Johnson, A., Chauncey, A. and Burkett, C., 2010. Self-regulated learning with MetaTutor: Advancing the science of learning with Metacognitive tools. In M. S. Khine & I. M. Saleh, eds. 2010. *New Science of Learning*. New York, NY: Springer. pp.225-247.

Baker, R.S., 2016. Stupid tutoring systems, intelligent humans. *International Journal of Artificial Intelligence in Education 26(2)*, pp. 600–614. doi: 10.1007/s40593-016-0105-0.

Baker, R.S., Corbett, A.T., Roll, I., Koedinger, K.R., Aleven, V., Cocea, M., Hershkovitz, A., de Carvalho, A.M.J.B., Mitrovic, A., Mathews, M., 2013. Modeling and Studying Gaming the System with Educational Data Mining. In Azevedo, R., & Aleven, V. (Eds.) *International Handbook of Metacognition and Learning Technologies.* pp. 97-116. New York, NY: Springer.

Bargiela, A. and Pedrycz, W., 2002. *Granular computing: an introduction.* Kluwer Academic Publishers, Boston, Dordrecht, London.

Biggs, J., 1998. Assessment and classroom learning: A role for summative assessment? *Assessment in Education: Principles, Policy & Practice 5(1)*, pp. 103–110. doi: 10.1080/0969595980050106.

Black, P. and Wiliam, D., 1998. Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice 5(1)*, pp.7-74.

Black, P. and Wiliam, D., 1998. *Inside the black box: raising standards through classroom assessment.* London: School of Education, King's College London.

Black, P. and Wiliam, D., 2009. Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education) 21(1)*, pp.5-31.

Black, P., Harrison, C., Hogden, J., Marshall, B. and Wiliam, D., 2005. Dissemination and evaluation: a response to Smith and Gorard. *Research Intelligence 93(7)*, p7.

Blanchard-Fields, F., 2007. Everyday problem solving and emotion: an adult developmental perspective. *Current Directions in Psychological Science 16 (1)*, pp.26–31.

Boud, D. and Falchikov, N., 1989. Quantitative studies of student self-assessment in higher education: a critical analysis of findings. *Higher education 18(5)*, pp.529-549.

Boud, D. and Molloy, E. (eds.), 2013. *Feedback in higher and professional education: Understanding it and doing it well*. New York, NY: Routledge.

Boud, D., 1995. *Enhancing learning through self-assessment*. Philadelphia: Routledge Falmer.

Boud, D., 2000. Sustainable assessment: rethinking assessment for the learning society. *Studies in continuing education 22(2)*, pp.151-167.

Boud, D., Keogh, R. and Walker, D., 1985. *Reflection: turning experience into learning*. Routledge.

Brusilovsky, P., Hsiao, I.H. and Folajimi, Y., 2011. QuizMap: open social student modeling and adaptive navigation support with TreeMaps. In *European Conference on Technology Enhanced Learning.* Springer Berlin Heidelberg. pp.71-82.

Bull, S., 2016. Negotiated learner modelling to maintain today's learner models. *Research and Practice in Technology Enhanced Learning 11(1)*, pp.1-29.

Bull, S. and Britland, M., 2007. Group interaction prompted by a simple assessed open learner model that can be optionally released to peers. In: Brusilovsky, P., Grigoriadou, M., Papanikolaou, K. (Eds.): Proceedings of *Workshop on Personalisation in E-Learning Environments at Individual and Group Level*, 11th International Conference on User Modeling, Corfu, Greece, 2007, pp.24-32

Bull, S., Johnson, M.D., Alotaibi, M., Byrne, W. and Cierniak, G., 2013. Visualising multiple data sources in an independent open learner model. In *International Conference on Artificial Intelligence in Education.* Springer Berlin Heidelberg. pp.199-208.

Bull, S., Johnson, M.D., Epp, C.D., Masci, D., Alotaibi, M. and Girard, S., 2014. Formative Assessment and Meaningful Learning Analytics. In proceedings of *International Conference on Advanced Learning Technologies,* pp.327-329.

Bull, S., Johnson, M.D., Masci, D. and Biel, C., 2015. Integrating and visualising diagnostic information for the benefit of learning. In: P. Reimann, S. Bull, and M Kickmeier-Rust (eds.) *Measuring and Visualizing learning in the information-rich classroom*. Abingdon, United Kingdom: Routledge. p.167.

Bull, S. and Kay, J., 2007. Student models that invite the learner in: the SMILI☺ open learner modelling framework. *International Journal of Artificial Intelligence in Education 17(2)*, pp.89-120.

Bull, S. and Kay, J., 2010. Open learner models. In *Nkambou, R., Mizoguchi, R. and Bourdeau, J. (eds.) Advances in Intelligent Tutoring Systems.* Berlin-Heidelberg: Springer. pp.301-322.

Bull, S. and Kay, J., 2013. Open learner models as drivers for Metacognitive processes. In: R. Azevedo & V. Aleven, (eds.) *International Handbook of Metacognition and Learning Technologies*. New York, NY: Springer, pp.349–365.

Bull, S. and 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., Mabbott, A. and Abu Issa, A.S., 2007. UMPTEEN: named and anonymous learner model access for instructors and peers. *International Journal of Artificial Intelligence in Education 17(3)*, pp.227-253.

Bull, S., McEvoy, A.T. and Reid, E., 2003. Learner models to promote reflection in combined desktop PC/mobile intelligent learning environments. In *Bull, S., Brna, P., Dimitrova, V. (eds.). Proceedings of the International Workshop on Learner Modelling for Reflection at 11th International Conference on Artificial Intelligent in Education*, pp.199-208.

Bull, S. and McKay, M., 2004. An open learner model for children and teachers: inspecting knowledge level of individuals and peers. In *International Conference on Intelligent Tutoring Systems.* Springer Berlin Heidelberg. pp.646-655.

Bull, S. and Pain, H., 1995. "Did I Say what I Think I Said, and Do You Agree with Me?": Inspecting and Questioning the Student Model. Proceedings of *Artificial Intelligence and Education, Association for the Advancement of Computing in Education*, Charlottesville, VA. pp.501-508.

Busch, L., 2014. Big Data, Big Questions. A Dozen Ways to Get Lost in Translation: Inherent Challenges in Large Scale Data Sets. *International Journal of Communication 8*, p.18.

Buse, R.P. and Zimmermann, T., 2012. Information needs for software development analytics. In *Proceedings of the 34th International Conference on Software Engineering*, Zurich, Switzerland, June 2-9, 2012. IEEE Press. pp.987-996.

Butler, D.L. and Winne, P.H., 1995. Feedback and self-regulated learning: a theoretical synthesis. *Review of Educational Research 65(3)*, pp.245-281.

Carey, M., Kriwaczek, F. and Ruger, S.M., 2000. A visualization interface for document searching and browsing. In *Proceedings of CIKM 2000 Workshop on New Paradigms in Information Visualization and Manipulation.* ACM Press. pp.24-28.

Cassidy, S., 2006. Learning style and student self-assessment skill. *Education and Training 48*(2/3). pp.170-177.

Chappuis, J., 2009. *Seven strategies of assessment for learning - a study guide.* Portland, OR: Pearson.

Chatti, M.A., Dyckhoff, A.L., Schroeder, U. and Thüs, H., 2012. A reference model for learning analytics. *International Journal of Technology Enhanced Learning 4(5-6)*, pp.318-331.

Chudowsky, N. and Pellegrino, J.W., 2003. Large-scale assessments that support learning: what will it take? *Theory into Practice 42(1)*, pp.75-83.

Crooks, T., 2011. Assessment for learning in the accountability era: New Zealand. *Studies in Educational Evaluation 37(1)*, pp.71-77.

Davis, D., Chen, G., Jivet, I., Hauff, C. and Houben, G.J., 2016. Encouraging metacognition and self-regulation in MOOCs through increased learner feedback. *Learning Analytics for Learners.* CEUR Vol. 1596, pp. 17-22.

Demmans Epp, C. and Bull, S., 2015. Uncertainty representation in visualizations of learning analytics for learners: current approaches and opportunities. *IEEE Transactions on Learning Technologies 8(3)*, pp.242-260.

Desautel, D., 2009. Becoming a thinking thinker: metacognition, self-reflection, and classroom practice. *Teachers College Record 111*(8), pp.1997-2020.

Dewey, J., 1933. *How we think: a restatement of the relation of reflective thinking to the educative process*. United States: Houghton Mifflin Harcourt (HMH).

Dimitrova, V. and Brna, P., 2016. From interactive open learner modelling to intelligent mentoring: STyLE-OLM and beyond. *International Journal of Artificial Intelligence in Education 26(1)* pp.332-349.

Dimitrova, V., 2003. STyLE-OLM: interactive open learner modelling. *International Journal of Artificial Intelligence in Education 13(1)*, pp.35-78.

Durall, E. and Gros, B., 2014. Learning analytics as a metacognitive tool. In *Proceedings of the 6th International Conference on Computer Supported Education CSEDU*, 1-3 April 2014, Barcelona, Spain. doi: 10.5220/0004933203800384. pp. 380-384

Duval, E., 2011. Attention please!: learning analytics for visualization and recommendation. In *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*, Banff, Alberta, Canada. ACM, pp.9-17.

Eysneck, M. and Keane, M., 2005. *Cognitive Psychology: A Student's Handbook.* Psychology Press.

Ferguson, R., 2012. Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning 4(5-6),* pp.304-317.

Flavell, J., & Wellman, H., 1977. Metamemory. In: R. V. Kail Jr. & J. Hagen (eds.) *Perspectives on the development of memory and cognition*. Hillsdale, NJ: Erlbaum. pp.3-33.

Flavell, J.H., 1976. Metacognitive aspects of problem solving. *The nature of intelligence 12*, pp.231-235.

Floridi, L., 2012. Big data and their epistemological challenge. *Philosophy & Technology*, pp.1-3.

Fonseca, F., Egenhofer, M., Davis, C. and Câmara, G., 2002. Semantic granularity in ontology-driven geographic information systems. *Annals of mathematics and artificial intelligence 36(1-2)*, pp.121-151.

Gabrielle, D.M., 2003. *Effects of Technology-Mediated Instructional Strategies on Motivation, Performance, and Self-Directed Learning*. PhD. Florida State University. Available at: http://diginole.lib.fsu.edu/islandora/object/fsu%3A182576 [Accessed 12.07.2016].

Ginon, B., Boscolo, C., Johnson, M.D. and Bull, S., 2016. Persuading an open learner model in the context of a university course: an exploratory study. In *International Conference on Intelligent Tutoring Systems*. Springer International Publishing. pp.307-313.

Gobert, J.D., Sao Pedro, M., Raziuddin, J. and Baker, R.S., 2013. From log files to assessment metrics: measuring students' science inquiry skills using educational data mining. *Journal of the Learning Sciences 22(4)*, pp.521-563.

Goldstein, I.P., 1982. The generic graph: a representation of the evolution of procedural knowledge. In *Intelligent Tutoring Systems.* Academic Press.

Good, R., 2011. Formative use of assessment information: it's a process, so let's say what we mean. *Practical Assessment, Research & Evaluation 16(3)*, p.2.

Goodman, B., Soller, A., Linton, F. and Gaimari, R., 1997. Encouraging student reflection and articulation using a learning companion. In *Proceedings of the AI-ED 97 World Conference on Artificial Intelligence in Education*. IOS Press. pp. 151-158.

Grow, G.O., 1991. Teaching learners to be self-directed. *Adult Education Quarterly 41(3)*, pp.125-149.

Handelsman, J., Ebert-May, D., Beichner, R., Bruns, P., Chang, A., DeHaan, R., Gentile, J., Lauffer, S., Stewart, J., Tilghman, S.M. and Wood, W.B., 2004. Scientific teaching. *Science 304*(5670), pp.521-522.

Hartley, D. and Mitrovic, A., 2001. *The effectiveness of open student modelling on learning*. Computer Science Department, University of Canterbury, New Zealand.

Hartman, H.J., 1998. Metacognition in teaching and learning: an introduction. *Instructional Science 26(1)*, pp.1-3.

Hattie, J. and Timperley, H., 2007. The power of feedback. *Review of Educational Research 77(1)*, pp.81-112.

Hattie, J.A., 2008. *Visible learning: a synthesis of over 800 meta-analyses relating to achievement.* London: Routledge.

Hattie, J.A. and Yates, G.C., 2014. Using feedback to promote learning. In V. A. Benassi, C. E. Overson, & C. M. Hakala (eds.) *Applying Science of Learning in Education: Infusing Psychological Science into the Curriculum*. Retrieved from the Society for the Teaching of Psychology: http://teachpsych.org/ebooks/asle2014/index.php, pp.45-58.

Hearst, M., 2006. Design recommendations for hierarchical faceted search interfaces. In *ACM SIGIR Workshop on Faceted Search,* Seattle, WA. ACM. pp.1-5.

Heer, J. and Agrawala, M., 2008. Design considerations for collaborative visual analytics. *Information visualization*, *7*(1), pp.49-62.

Heritage, M., Kim, J., Vendlinski, T. and Herman, J., 2009. From evidence to action: A seamless process in formative assessment?. *Educational Measurement: Issues and Practice*, *28*(3), pp.24-31. Wiley.

Ikeda, R., Cho, J., Fang, C., Salihoglu, S., Torikai, S. and Widom, J., 2012. *Provenance-based Debugging and drill-down in data-oriented Workflows.* Available at: http://ilpubs.stanford.edu:8090/1008/ (Accessed: 12 December 2016).

Jackson, C.W. and Larkin, M.J., 2002. Rubric: Teaching students to use grading rubrics. *Teaching exceptional children*, *35*(1), p.40.

Jacoby, J., 1977. Information load and decision quality: Some contested issues. *Journal of Marketing Research*, pp.569-573.

Jern, M., 1997. Information drill-down using web tools. In *Visualization in Scientific Computing'97* (pp. 9-20). Springer Vienna.

Johan, R. and Bull, S., 2009. Consultation of Misconceptions Representations by Students in Education-Related Courses. In C.K. Looi, G. McCalla and B. Bredeweg (eds.) 2005. Artificial intelligence in education: Supporting learning through intelligent and socially informed technology. Amsterdam: IOS Press. pp. 565-572.

Johnson, M.D. and Bull, S., 2009. Belief exploration in a multiple-media open learner model for basic harmony. *Artificial Intelligence in Education: Building Learning Systems that Care: From Knowledge Representation to Affective Modelling*, *200*, p.299.

Johnson, M.D. and Bull, S., 2011. Optional finer granularity in an open learner model. In *International Conference on Artificial Intelligence in Education* (pp. 478-480). Springer Berlin Heidelberg.

Johnson, M.D., Bull, S. and Kickmeier-Rust, M., 2013a. Student competency visualisation for teachers. In *3rd Int. Workshop on Teaching Analytics, EC-TEL 2013*. CEUR.

Johnson, M.D., Cierniak, G., Hansen, C., Bull, S., Wasson, B., Biel, C. and Debus, K., 2013b. Teacher approaches to adopting a competency based open learner model. In *Proceedings of the 21st International Conference on Computers in Education.* Indonesia: Asia-Pacific Society for Computers in Education.

Johnson, M.D., Ginon, B., Masci, D., Baber, C., Bull, S., 2015. Deliverable D4.3 – second release of visualisation and OLM services and tools. http://css-kmi.tugraz.at/mkrwww/leas-box/downloads/D4.3.pdf [accessed 30/08/2016]

Johnson, M.D., Reimann, P., Bull, S. and Fujita, N., 2011. Epistemic beliefs and open learner models. In *Proceedings of the International Workshop on Personalization Approaches in Learning Environments*. CEUR Vol. 732, pp.19-24.

Kalz, M., 2015. Lifelong Learning and its support with new technologies. In J. D. Wright (ed.) *International Encyclopedia of the Social & Behavioral Science 14(2).* Oxford: Elsevier. pp.93-99.

Kaupp, J., Frank, B. and Watts, C., 2013. Evaluation of software tools supporting outcomes-based continuous program improvement processes. In *Proceedings of the Canadian Engineering Education Association*, Montreal, QC.

Kay, J. and Bull, S., 2015. New opportunities with open learner models and visual learning analytics. In *Proceedings of 17th International Conference on Artificial Intelligence in Education,* Madrid, Spain. Springer International Publishing. pp.666-669.

Kay, J. and Lum, A., 2005. Exploiting readily available web data for scrutable student models. In C.K. Looi, G. McCalla and B. Bredeweg (eds.) *Artificial intelligence in education: Supporting learning through intelligent and socially informed technology*. Amsterdam: IOS Press. pp.338-345.

Kay, J., 1994. The UM toolkit for cooperative user modelling. *User Modeling and User-Adapted Interaction 4(3)*, pp.149-196.

Kerly, A., Ellis, R. and Bull, S., 2008. CALMsystem: a conversational agent for learner modelling. *Knowledge-Based Systems 21*(3), pp.238-246.

Khan, M., Hussain, F. and Khan, I., 2014. Single level drill down interactive visualization technique for descriptive data mining results. *International Journal of Grid and Distributed Computing 7*(4), pp.33-40.

Kickmeier-Rust, M. and Albert, D., 2016. Theory-driven learning analytics and open learner modelling: the teacher's toolbox of tomorrow?. In *Proceedings of the 6th Workshop on Personalization Approaches in Learning Environments, 24th Conference on User Modeling, Adaptation, and Personalization*, CEUR workshop proceedings.

Knowles, M.S. and Knowles, K., 1975. *Self-directed learning: a guide for learners and teachers*. New York: Cambridge, the Adult Education Co.

Kolb, D.A., 1976. *Learning styles inventory*. Boston.

Kump, B., Seifert, C., Beham, G., Lindstaedt, S.N. and Ley, T., 2012. Seeing what the system thinks you know: visualizing evidence in an open learner model. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*. ACM. pp.153-157.

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(3)*, pp.255-271.

Lengler, R. and Eppler, M.J., 2007. Towards a periodic table of visualization methods for management. In *IASTED Proceedings of the Conference on Graphics and Visualization in Engineering (GVE 2007), Clearwater, Florida, USA*.

Li, L. and Kay, J., 2005. Assess: promoting learner reflection in student self-assessment. In *Proceedings of the Workshop on Learner Modelling for Reflection, to Support Learner Control, Metacognition and Improved Communication between Teachers and Learners (LeMoRe05)*, Series Assess: Promoting Learner Reflection in Student Self-Assessment. pp.32-41.

Liang, M., Guerra, J., Marai, G.E. and Brusilovsky, P., 2012. Collaborative e-learning through open social student modeling and progressive zoom navigation. In *Proceedings of 8th International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom), 2012,* Pennsylvan, USA. IEEE. pp.252-261.

Livingston, J.A., 1996. Effects of metacognitive instruction on strategy use of college students. Unpublished manuscript. State University of New York at Buffalo.

Lloyd, T. and Bull, S., 2006. A haptic learner model. *International Journal of Continuing Engineering Education and Life Long Learning 16(1-2)*, pp.137-149.

Long, Y. and Aleven, V., 2013. Supporting students' self-regulated learning with an open learner model in a linear equation tutor. In *Proceedings of International Conference on Artificial Intelligence in Education*. Springer Berlin Heidelberg. pp.219-228.

Luecht, R.M., 2013. An introduction to assessment engineering for automatic item generation. *Automatic Item Generation: Theory and Practice*, pp.59-76.

Mabbott, A. and Bull, S., 2004. Alternative views on knowledge: presentation of open learner models. In *International Conference on Intelligent Tutoring Systems.* Springer Berlin Heidelberg. pp.689-698.

Mabbott, A. and Bull, S., 2006. Student preferences for editing, persuading, and negotiating the open learner model. In *International Conference on Intelligent Tutoring Systems*. Springer Berlin Heidelberg. pp.481-490.

Mabbott, A. and Bull, S., 2006. Student preferences for editing, persuading, and negotiating the open learner model. In *International Conference on Intelligent Tutoring Systems*. Springer Berlin Heidelberg. pp.481-490.

Mabbott, A., 2009. *User choice in viewing and interacting with open learner models.* University of Birmingham.

Mahdavi, M., 2014. An overview: metacognition in education. *International Journal of Multidisciplinary and Current Research*, pp.529-535.

Markauskaite, L. and Jacobson, M.J., 2015. Tracking and assessing students' learning strategies in model-based learning environments. In: P. Reimann, S. Bull, and M Kickmeier-Rust, (eds.) *Measuring and Visualizing learning in the Information-Rich Classroom*. Abingdon, United Kingdom: Routledge. p.137.

Mathews, M., Mitrovic, A., Lin, B., Holland, J. and Churcher, N., 2012. Do your eyes give it away? Using eye tracking data to understand students' attitudes towards open student model representations. In *International Conference on Intelligent Tutoring Systems*. Springer Berlin Heidelberg. pp.422-427.

Mayer, R.E., 2013. Problem Solving. In Reisberg, D. (Ed.) *The Oxford Handbook of Cognitive Psychology*. OUP.

McCalla, G., 2004. The ecological approach to the design of e-learning environments: purpose-based capture and use of information about learners. *Journal of Interactive Media in Education 1*.

McCalla, G.I. and Greer, J.E., 1994. Granularity-based reasoning and belief revision in student models. In *Student Modelling: The Key to Individualized Knowledge-Based Instruction*. Springer Berlin Heidelberg. pp.39-62.

Mello, L.V., 2016. Fostering postgraduate student engagement: online resources supporting self-directed learning in a diverse cohort. *Research in Learning Technology 24*.

Mitrovic, A. and 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.

Mitrovic, A. and Ohlsson, S., 2016. Implementing CBM: SQL-Tutor after fifteen years. *International Journal of Artificial Intelligence in Education 26(1)*, pp.150-159.

Mitrovic, A. and Weerasinghe, A., 2009. Revisiting ill-definedness and the consequences for ITSs. *Artificial Intelligence in Education: Building Learning Systems that Care from Knowledge Representation to Affective Modelling*, pp.375-382.

Morales, R., Van Labeke, N., Brna, P. and Chan, M.E., 2008. Open learner modelling as the keystone of the next generation of adaptive learning environments. In C. Mourlas and P Germanakos (eds.) *Intelligent User Interfaces: Adaptation and Personalization Systems and Technologies*. United States: Information Science Reference. pp.288-312.

Nicol, D., 2009. Assessment for learner self-regulation: enhancing achievement in the first year using learning technologies. *Assessment and Evaluation in Higher Education 34(3)*, pp.335-352.

Nicol, D.J. and Macfarlane-Dick, D., 2006. Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education 31(2)*, pp.199-218.

Norman, K.L., 1991. *The psychology of menu selection: designing cognitive control at the human-computer interface*. Intellect Books.

Pachler, N., Daly, C., Mor, Y. and Mellar, H., 2010. Formative e-assessment: practitioner cases. *Computers & Education 54(3)*, pp.715-721.

Paiva, A., Self, J. and Hartley, R., 1995. Externalising learner models. In *Proceedings of World Conference on Artificial Intelligence in Education*. pp.509-516.

Papanikolaou, K.A., Grigoriadou, M., Kornilakis, H. and Magoulas, G.D., 2003. Personalizing the interaction in a web-based educational hypermedia system: the case of INSPIRE. *User Modeling and User-Adapted Interaction 13(3)*, pp.213-267.

Papies, E. K. and Aarts, H., 2016. Automatic self-regulation: from habit to goal pursuit. In: Vohs, K. D. and Baumeister, R. F. (eds.) *Handbook of Self-Regulation: Research, Theory, and Applications.* Guilford Press: New York. ISBN 9781462520459.

Paris, S.G. and Winograd, P., 1990. How metacognition can promote academic learning and instruction. *Dimensions of Thinking and Cognitive Instruction 1*, pp.15-51.

Pedrycz, W., 2014. Allocation of information granularity in optimization and decision-making models: towards building the foundations of granular computing. *European Journal of Operational Research 232(1)*, pp.137-145.

Pellegrino, J.W., Chudowsky, N. and Glaser, R., 2001. *Knowing what students know: the science and design of educational assessment*. Washington, DC: National Academies Press.

Pellegrino, J.W., Wilson, M.R. and Koenig, J.A., 2014. *Developing assessments for the next generation science standards.* Washington, DC: National Academies Press.

Perez-Marín, D.R., 2007. *Adaptive computer assisted assessment of free-text students' answers: an approach to automatically generate students' conceptual models*. Doctoral dissertation, Universidad Autónoma de Madrid.

Pinchok, N. and Brandt, W.C., 2009. Connecting formative assessment research to practice: an introductory guide for educators. *Learning Point Associates.* Available at: http://files.eric.ed.gov/fulltext/ED509943.pdf [Accessed 12.07.2016].

Pintrich, P.R., 2000. The role of goal orientation in self-regulated learning. In: M. Boekaerts, P.R. Pintrich, and M. Zeidner (eds.) *Handbook of Self-Regulation*. San Diego, CA: Elsevier Science. pp.451-502.

Pintrich, P.R., 2003. Motivation and classroom learning. *Handbook of Psychology* doi: 10.1002/0471264385.wei0706

Popham, J.W., 2011. *Transformative assessment in action: an inside look at applying the process*. Alexandria, VA: Association for Supervision & Curriculum Development.

Price, V., 2015. Exploring effectiveness and rationale of different assessment types. *Journal of Initial Teacher Inquiry 1*, p.12.

Ras, E., Whitelock, D. and Kalz, M., 2015. The promise and potential of e-assessment for learning. In P. Reimann, S. Bull, and M Kickmeier-Rust (eds.) *Measuring and Visualizing Learning in the Information-Rich Classroom*. Routledge. p.21.

Raybourn, E. and Regan, D., 2011. Exploring e-portfolios and independent open learner models: toward army learning concept. In *Interservice/Industry Training, Simulation, and Education Conference Proceedings*.

Redecker, C. and Johannessen, Ø., 2013. Changing assessment—towards a new assessment paradigm using ICT. *European Journal of Education 48(1)*, pp.79-96.

Reimann, P., Halb, W., Bull, S. and Johnson, M., 2011. Design of a computer-assisted assessment system for classroom formative assessment. In *Proceedings of 14th International Conference on Interactive Collaborative Learning (ICL).* IEEE. pp.465-472.

Rosenfeld, L. and Morville, P., 2006. *Information architecture for the world wide web.* 2nd edition. Sebastopol, CA, United States: O'Reilly Media Inc., USA.

Scardamalia, M., 2002. Collective cognitive responsibility for the advancement of knowledge. *Liberal Education in a Knowledge Society 97*, pp.67-98.

Schoenfeld, A.H., 1987. What's all the fuss about metacognition. *Cognitive Science and Mathematics Education 189*.

Schon, D.A. and DeSanctis, V., 1986. The reflective practitioner: how professionals think in action. *The Journal of Continuing Higher Education 34(3)*, pp.29-30.

Schraw, G. and Dennison, R.S., 1994. Assessing metacognitive awareness. *Contemporary Educational Psychology 19(4)*, pp.460-475.

Self, J., 1998. The defining characteristics of intelligent tutoring systems research: ITSs care, precisely. *International Journal of Artificial Intelligence in Education (IJAIED) 10*, pp.350-364.

Self, J.A., 1990. Bypassing the intractable problem of student modelling. *Intelligent Tutoring Systems: At the Crossroads of Artificial Intelligence and Education 41*, pp.1-26.

Sharpe, R., Benfield, G. and Francis, R., 2006. Implementing a university e-learning strategy: levers for change within academic schools. *Research in Learning Technology 14(2)*. Routledge.

Shepard, L.A., 2005. Formative assessment: caveat emptor. ETS Invitational Conference 2005. The Future of Assessment: Shaping Teaching and Learning, New York. Retrieved from http://www.cpre.org/ccii/images/stories/ccii_pdfs/shepard_formative assessment_caveat_emptor.pdf [Accessed 25-05-2016].

Shneiderman, B., 1996. The eyes have it: a task by data type taxonomy for information visualizations. In *Proceedings of IEEE Symposium on Visual Languages.* IEEE. pp.336-343.

Shute, V.J., 2008. Focus on formative feedback. *Review of Educational Research 78*(1), pp.153-189.

Siemens, G., 2010. What are learning analytics? Available at: http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics/ [Accessed: 10-03-2011].

Silver, E. A., 1985. *Teaching and learning mathematical problem solving: multiple research perspectives.* New York, NY, United States: Lawrence Erlbaum Associates.

Sison, R. and Shimura, M., 1998. Student modeling and machine learning. *International Journal of Artificial Intelligence in Education (IJAIED) 9*, pp.128-158.

Smyth, M., Collins, A., Morris, P. and Levy, P., 1994. *Cognition in action.* Psychology Press Ltd.

Stiggins, R., Arter, J. Chappuis, S., and Chappuis, J., 2012. *Classroom assessment for student learning*. Upper Saddle River, NJ: Pearson Education.

Tanner, H., and Jones, S., 1999. Scaffolding metacognition: reflective discourse and the development of mathematical thinking. Proceedings of the *British Educational Research Association Conference*.

Tervakari, A.M., Silius, K., Koro, J., Paukkeri, J. and Pirttilä, O., 2014. Usefulness of information visualizations based on educational data. In *2014 IEEE Global Engineering Education Conference (EDUCON), 3-5 April 2014, Istanbul, Turkey*. pp.142-151.

Thompson, K. and Reimann, P., 2010. Patterns of use of an agent-based model and a system dynamics model: The application of patterns of use and the impacts on learning outcomes. *Computers & Education 54(2)*, pp.392-403.

Tominski, C., Abello, J., van Ham, F., and Schumann, H., 2006. Fisheye tree views and lenses for graph visualization. In: *10th International Conference on Information Visualisation*, London, UK. pp. 17-24. doi:10.1109/IV.2006.54

Uther, J. and Kay, J., 2003. VlUM, a web-based visualisation of large user models. In *International Conference on User Modeling*. Springer Berlin Heidelberg. pp.198-202.

van Aalst, J., Mu, J. and Yang, Y., 2015. Formative assessment of computer-supported collaborative learning and knowledge building. In: P. Reimann, S. Bull, and M Kickmeier-Rust, (eds.) *Measuring and Visualizing Learning in the Information-Rich Classroom*. Routledge. p.154.

van Labeke, N., Brna, P. and Morales, R., 2007. Opening up the interpretation process in an open learner model. *International Journal of Artificial Intelligence in Education 17(3)*, pp.305-338.

Vatrapu, R., Reimann, P., Bull, S. and Johnson, M., 2013. An eye-tracking study of notational, informational, and emotional aspects of learning analytics representations. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*. New York, NY: ACM. pp. 125-134.

Veenman, M.V., Van Hout-Wolters, B.H. and Afflerbach, P., 2006. Metacognition and learning: conceptual and methodological considerations. *Metacognition and Learning 1(1)*, pp.3-14.

Veenman, M.V., Wilhelm, P. and Beishuizen, J.J., 2004. The relation between intellectual and metacognitive skills from a developmental perspective. *Learning and Instruction 14(1)*, pp.89-109.

Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos, J. L., 2013. Learning analytics dashboard applications. *American Behavioral Scientist* 57(10), pp.1500–1509. doi:10.1177/0002764213479363

Wang, Y., and Chiew, V., 2010. On the cognitive process of human problem solving. *Cognitive Systems Research 11(1)*, pp.81-92.

Wallace, D.F., Anderson, N.S. and Shneiderman, B., 1987. Time stress effects on two menu selection systems. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 31(7)*, SAGE Publications. pp.727–731. doi: 10.1177/154193128703100708.

Weber, G. and Brusilovsky, P., 2001. ELM-ART: An adaptive versatile system for web-based instruction. *International Journal of Artificial Intelligence in Education (IJAIED) 12*, pp.351-384.

Weinert, F.E., 1987. Metacognition and motivation as determinants of effective learning and understanding. In *Metacognition, Motivation, and Understanding*, Hillsdale, NJ: Lawrence Erlbaum, pp.1–16.

Wellman, H., 1985. The origins of metacognition. In D.L. Forrest-Pressley, G.E. MacKinnon, & G.T. Waller (eds.) *Metacognition, Cognition and Human Performance: Theoretical Performance 1*. Orlando, FL: Academic Press. pp.1-31.

Whitelock, D., 2010. Activating assessment for learning: are we on the way with web 2.0. *Web* 2, pp.319-342.

Winne, P.H. and Hadwin, A.F., 2008. The weave of motivation and self-regulated learning. In: D.H. Schunk & B.J. Zimmerman (eds.) *Motivation and Self-Regulated Learning: Theory, Research, and Applications.* New York, NY: Lawrence Erlbaum Associates. pp.297-314.

Woolf, B. P., 2008. *Building intelligent interactive tutors: student-centered strategies for revolutionizing e-learning.* Amsterdam: Morgan Kaufmann Publishers/Elsevier.

Yancey, K.B., 1998. *Reflection in the writing classroom*. United States: Utah State University Press.

Yi, J.S., ah Kang, Y., Stasko, J. and Jacko, J., 2007. Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on Visualization and Computer Graphics 13(6)*, pp.1224-1231.

Zapata-Rivera, D., Hansen, E., Shute, V.J., Underwood, J.S. and Bauer, M., 2007. Evidence-based approach to interacting with open student models. *International Journal of Artificial Intelligence in Education 17(3)*, pp.273-303.

Zapata-Rivera, J.D. and Greer, J.E., 2004. Interacting with inspectable Bayesian student models. *International Journal of Artificial Intelligence in Education 14(2)*, pp.127-163.

Zohar, A. and Dori, Y.J., 2003. Higher order thinking skills and low-achieving students: are they mutually exclusive? *The Journal of the Learning Sciences 12(2)*, pp.145-181.

APPENDIX 1: PARTICIPANT CONSENT FORM

APPENDIX 2: PARTICIPANT INFORMATION SHEET

Thank you for taking the time to participate in the study.
Below is some further information about its research content
and how your information will be used.

The Study

The technology you have helped evaluate is an educational system which aims to direct learners' attention towards their knowledge and understanding. The system provides an opportunity for learners to reflect on their current understanding, difficulties and misconceptions, rather than providing a conventional test or exercise, with a score being given upon its completion. The system considers recent responses to multiple choice questions to infer beliefs held by you as the learner. The role of the system is to help learners make appropriate decisions in how they approach their learning.

The technology has been developed to investigate learners' use of different levels of detail in open learner model presentation, with particular reference to self-assessment accuracy and the completion of different tasks during interaction. It is theorised that access to open learner model information increases the accuracy of self-assessment and will allow learners to appropriately focus their learning towards problems identified in their knowledge. The outcome of the research will help identify implications for educational environments of this type.

Use of Information

Interaction logs will be used to re-create interaction patterns and to provide quantitative data indicating the state of the learner model and a model of your self-assessment at the point open learner model features are accessed. This is to elicit the context(s) of use for specific features. Survey data and any open ended comments you provide will be compared against interaction logs to provide a fuller picture of system usage.

Data collected is stored in a secure location on file servers and in locked filing cabinets at the University of Birmingham. Only the researcher/supervisor can access the information. Participants will not be identified by name in any analysis, results or subsequent publication of findings; data is stored, accessed and referred to by sample number and is treated confidentially.

Researchers' Contact Information

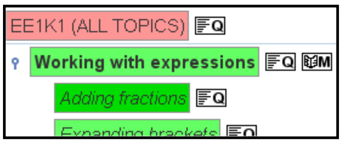
If you would like more information or wish to withdraw permission allowing your data to be used in research, please contact the researcher or supervisor (details below). Permission may be withdrawn at any point up until publication of the findings.

APPENDIX 3: POST-USAGE SURVEY

Thinking about your use of the OLM, please respond to each of the following question, indicating your agreement with each statement on a scale of 5 (strongly agree) to 1 (strongly disagree).

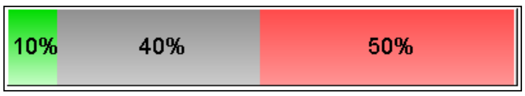
	Strongly Agree 5	4	3	2	Strongly Disagree 1
A. General					
1) I am good at self-assessment					
2) *I viewed my learner model…*					
a) to see general progress					
b) for information, but wasn't sure as to what					
c) for information about the last question answered					
d) for information about a previous question answered					
e) for information about a question about to be attempted					

B. List of Topics and Concepts



	5	4	3	2	1
3) I understood the list presentation of my learner model					
4) The list presentation of my learner model was useful					
5) *The list presentation of my learner model was useful…*					
a) to help me plan what to do next					
b) when identifying what I understood					
c) when identifying problems					
d) to help me think about how I am approaching my learning					

C. Skill Meter



	5	4	3	2	1
6) I understood the skill meter presentation of my learner model					
7) The skill meter presentation of my learner model was useful					
8) *The skill meter presentation of my learner model was useful…*					
a) to help me plan what to do next					
b) when identifying what I understood					
c) when identifying problems					
d) to help me think about how I am approaching my learning					

D. My Beliefs

$\log_b x^n = \log_b \left(\frac{x}{n}\right)$ The effect of an exponent in a logarithm value (n) is the same as dividing the value (x) by the exponent (n). This is the simpler form.

	Strongly Agree 5	4	3	2	Strongly Disagree 1
9) I understood the ‘my beliefs’ presentation of my learner model					
10) The ‘my beliefs’ presentation of my learner model was useful					
11) *The ‘my beliefs’ presentation of my learner model was useful…*					
a) to help me plan what to do next					
b) when identifying what I understood					
c) when identifying problems					
d) to help me think about how I am approaching my learning					

E. Expert Beliefs

$\log_b pq = \log_b p + \log_b q$ The logarithm of two values multiplied together is the same as the logarithm of the first plus the logarithm of the second. This is the simpler form.

12) I understood the ‘expert beliefs’ presentation of my learner model					
13) The ‘expert beliefs’ presentation of my learner model was useful					
14) *The ‘expert beliefs’ presentation of my learner model was useful…*					
a) to help me plan what to do next					
b) when identifying what I understood					
c) when identifying problems					
d) to help me think about how I am approaching my learning					

F. Drill Down Arrangement of the Learner Model

My learner model allowed me to view more detailed information by clicking on summaries of my understanding (e.g. clicking the list of topics and concepts to view the skill meter or clicking the skill meter to view my beliefs). This is a ‘drill down’ arrangement of the learner model.

15) I understood the ‘drill down’ arrangement of the learner model					
16) The ‘drill down’ arrangement of my learner model was useful					

Admin:

Sample Number: ______________________________

APPENDIX 4: INTRODUCTION AT START OF EVALUATIONS

Before Session Checklist

- Have enough copies of consent forms, debrief forms, surveys.
- Technology is running without issue and self-registration service is operational.
- Course lecturers have contacted participants to make them aware of the session.

Introduction to Technology

- Verbally confirm the domain content the technology and where this fits in to their university programme.
- Log in to test account, which already has a partially complete model. Explain that participants will need to complete self-registration. Email will be automatically sent with confirmation details. Researcher is on hand for any problems.
- Cover aspects of the formative purpose of the technology – intended to show current state of understanding on an area of the domain. It is not a test. It is not summative. It does not count towards any part of the module mark. Etc. (Avoid saying anything about educational benefit.)
- Show how to select an area of the domain from the overview list on which to answer questions.
- Answer sample question. Explain multiple choice question concept, how to answer and roughly what the system will do to validate the question response.
- Show how to submit a question and explain the concept of the self-assessment bar at the bottom which is used to submit the question.
- Explain the questioning cycle concept and that participant will have to break away from this to change the focus and to look at the learner model. Explain that the model may be inspect at any point, including during answering a question. The same question will be returned to.
- Switch to viewing the OLM. Explain the high level overview (list of topic and concepts), what the different colours mean, and how to hover over and view label summaries of what is shown if required.
- Explain that to get at more detailed information required to click on content (to drill into the model). Select a random item to do this for. Do this at least twice, selecting both correct and incorrect content as not to bias towards a particular information category.
- This will present a skill meter. Explain the different information categories in the skill meter and how it relates. Explain how the model weights more recent information to greater extent, hence different weightings. Indicate that from there they may switch to viewing a different concept, go back to questions, or can drill deeper to see specific beliefs that the model has inferred.
- Show the beliefs section for the chosen concept. Show that this is available in two different forms, text and something domain specific (e.g. maths notation) and explain how they describe the same thing. Show that they may also compare this to the domain (what an expert would have in their model) and how to make direct comparison.
- Answer any specific questions. Questions may be asked throughout the session(s) if unsure.
- Explain the format of the lab session, independent working.
- Ma(I) and Bo(I) will have to answer a series of questions before viewing the model. Make this clear and draw attention to the progress bar to see how far through the questions they are.

- Remind participants that they can stop at any time and there is no obligation to participate and can use it anyway even if they do not want to be part of a study. At the end they will be asked if they asked if some data can be used for research and there will be a survey.

During Session

- Be on hand to answer any queries relating to the operation of the technology etc.

End of Session Checklist (Lab Studies)

- Hand out survey and consent form.
- When survey and consent form complete, collect these and hand out debrief sheet.
- Verbally reiterate right of withdrawal and contact information.
- Make sure matching sample number is added to the consent form and the survey.

End of Session Checklist (8 Week Study)

- Nothing at end of first session.

End of Term Checklist (8 Week Study)

- As per the end of session check list.

APPENDIX 5: IDENTIFICATION OF POINTS OF MEASUREMENT BETWEEN BEHAVIOURAL AND DIAGNOSTIC MODELS

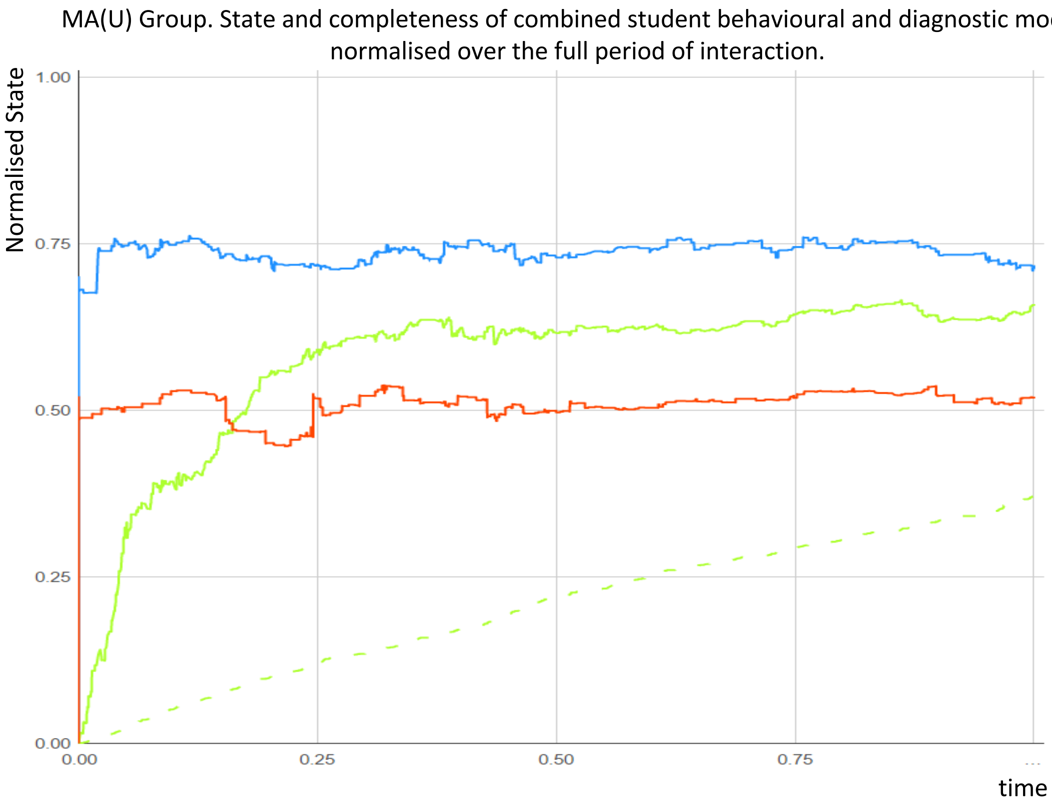
To allow for comparison between behavioural and diagnostic models of student understanding, the two models are each kept to the same precision. The models are compared where the analysis requires a calculation of student self-assessment accuracy, and this is taken as the mean discrepancy between the two models, per participant. To indicate changes across time measurement points at the start and end of interaction are required, however, in 3 of the 5 experimental groups the models are initially empty and therefore there is a mismatch in the accuracy between these two points in time. This section identifies suitable measurement points for the self-assessment accuracy calculation.

1. Points of Measurement (Study 1)

For each of the 4 experimental groups, Figure 32 to Figure 35 show the values for the average model state (blue line), the average self-assessment value (red line), and level of completeness of the model for those concepts that are active (green line). Time on the x-axis is normalised between the first OLM interaction event (0.0) and the last interaction event (1.0) to make participants comparable.

It can be seen in the MA(I) and BO(I) groups where the model already has information this is already at a good level of completeness at time 0.0, thus is taken as the initial point. For the two groups where the model is initially empty, this starts to plateau at around the end of the first quarter, thus time=0.25 is taken as the initial point. The dotted green lines show the overall absolute completeness of the learner model, taking into consideration concepts for which there is no information. It is noteworthy that where the model is instantiated initially, there is a good level of the domain covered, whereas when the model is built over time it does not reach similar levels at any point during interaction. It may be considered that the accuracy of the inferred state of the learner is to a lesser precision.

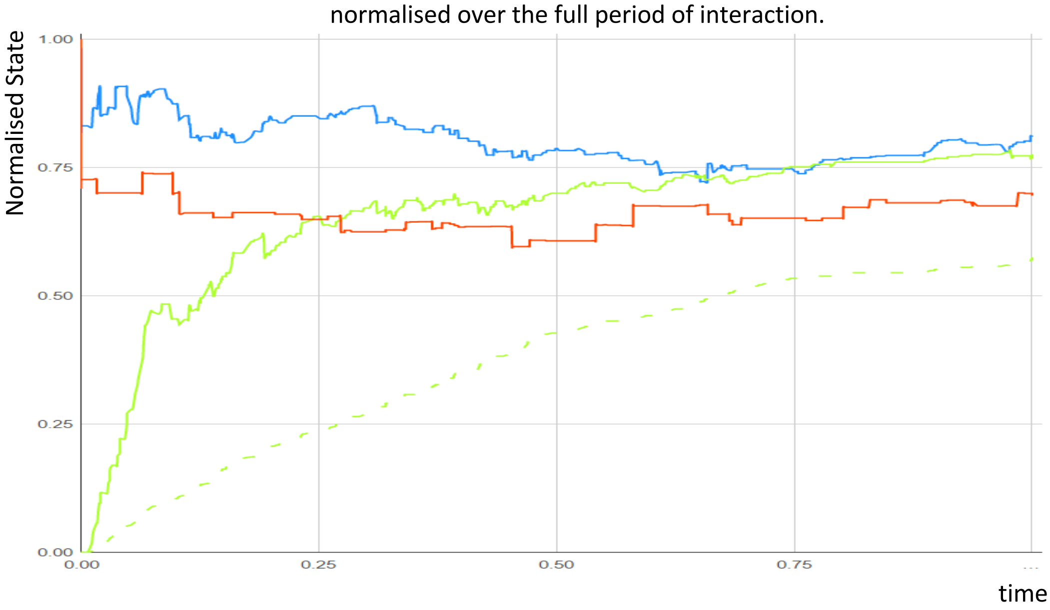
All four groups may be used to calculate change in self-assessment accuracy, however the two uninstantiated groups are to a lesser precision than the instantiated ones, and due to the temporal overhead of the period of model initialisation, the time period for change is less for the instantiated groups. Both therefore have limitations in the information provided.



Key:
Blue – behavioural - learner model state for **knowledge** (0.0 weak, 1.0 strong)
Red – diagnostic - model of the student's self-assessed **confidence** in their ability (0.0 weak, 1.0 strong)
Solid Green – **completeness** of the model (**active concepts**) (0.0 incomplete, 1.0 fully complete)
Dotted Green – **completeness** of the model (**absolute**) (0.0 incomplete, 1.0 fully complete)

Figure 32: MA(U) state and completeness of models over time.

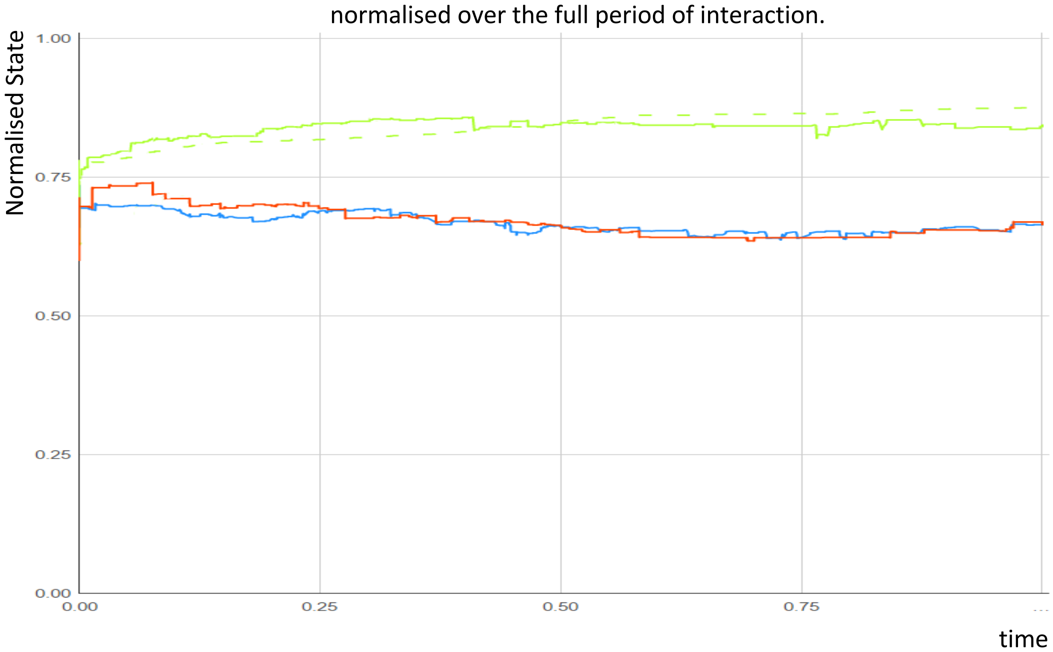
BO(U) Group. State and completeness of combined student behavioural and diagnostic models normalised over the full period of interaction.



Key:
Blue – behavioural - learner model state for **knowledge** (0.0 weak, 1.0 strong)
Red – diagnostic - model of the student's self-assessed **confidence** in their ability (0.0 weak, 1.0 strong)
Solid Green – **completeness** of the model (**active concepts**) (0.0 incomplete, 1.0 fully complete)
Dotted Green – **completeness** of the model (**absolute**) (0.0 incomplete, 1.0 fully complete)

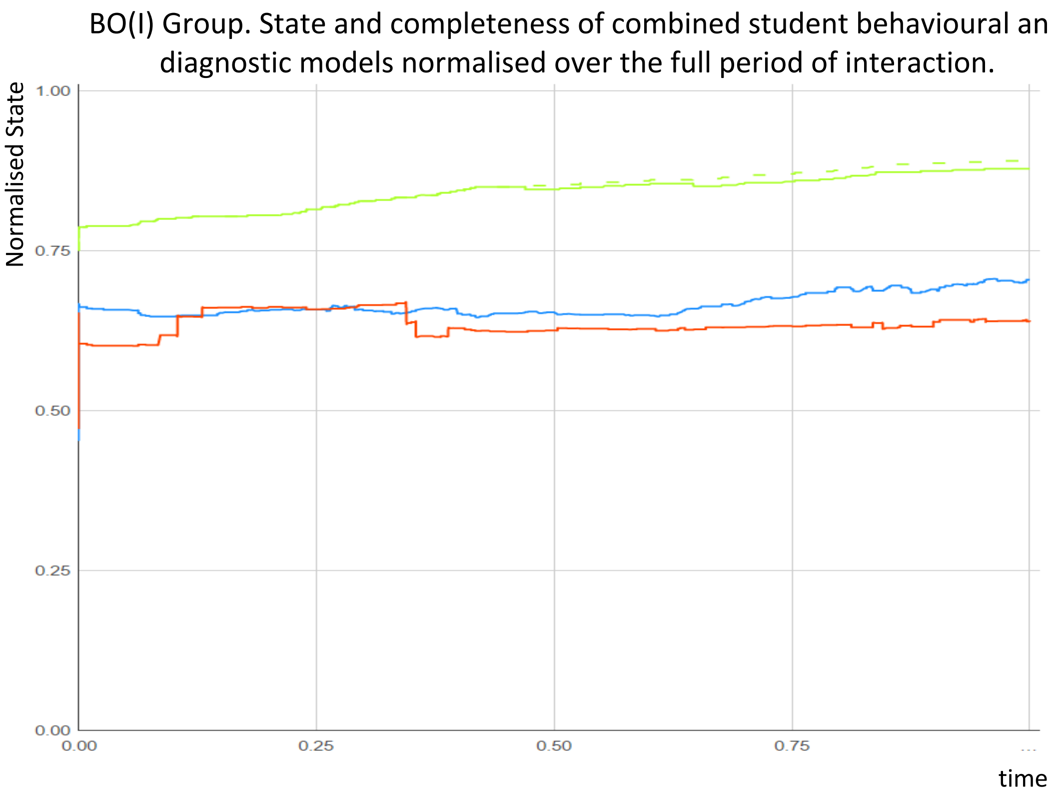
Figure 33: BO(U) state and completeness of models over time.

MA(I) Group. State and completeness of combined student behavioural and diagnostic models normalised over the full period of interaction.



Key:
Blue – behavioural - learner model state for **knowledge** (0.0 weak, 1.0 strong)
Red – diagnostic - model of the student's self-assessed **confidence** in their ability (0.0 weak, 1.0 strong)
Solid Green – **completeness** of the model (**active concepts**) (0.0 incomplete, 1.0 fully complete)
Dotted Green – **completeness** of the model (**absolute**) (0.0 incomplete, 1.0 fully complete)

Figure 34: MA(I) state and completeness of models over time.



Key:
Blue – behavioural - learner model state for **knowledge** (0.0 weak, 1.0 strong)
Red – diagnostic - model of the student's self-assessed **confidence** in their ability (0.0 weak, 1.0 strong)
Solid Green – **completeness** of the model (**active concepts**) (0.0 incomplete, 1.0 fully complete)
Dotted Green – **completeness** of the model (**absolute**) (0.0 incomplete, 1.0 fully complete)

Figure 35: BO(I) state and completeness of models over time.

In considering the average discrepancies between the learner model state (behavioural) and model of student self-assessment (diagnostic) it can be seen that for those who completed a series of questions beforehand, student self-assessment, as an average is already fairly accurate, whereas there is a discernible discrepancy in the case of the initially empty models. However, as mentioned above, these are with greatly differing amounts of data in the models, making it difficult to draw comparison between instantiated and instantiated cases.

In considering the two groups that have mostly complete models, BO(I) shows a divergence, with the diagnostic model starting to rise latent to the behavioural model. This is potentially the case as it may take participants time to build trust, following a change in cognitive state (e.g. something new has been learnt). For MA(I) there is a minimal decrease in the behavioural model, but with the behavioural model aligning with this over the period of interaction.

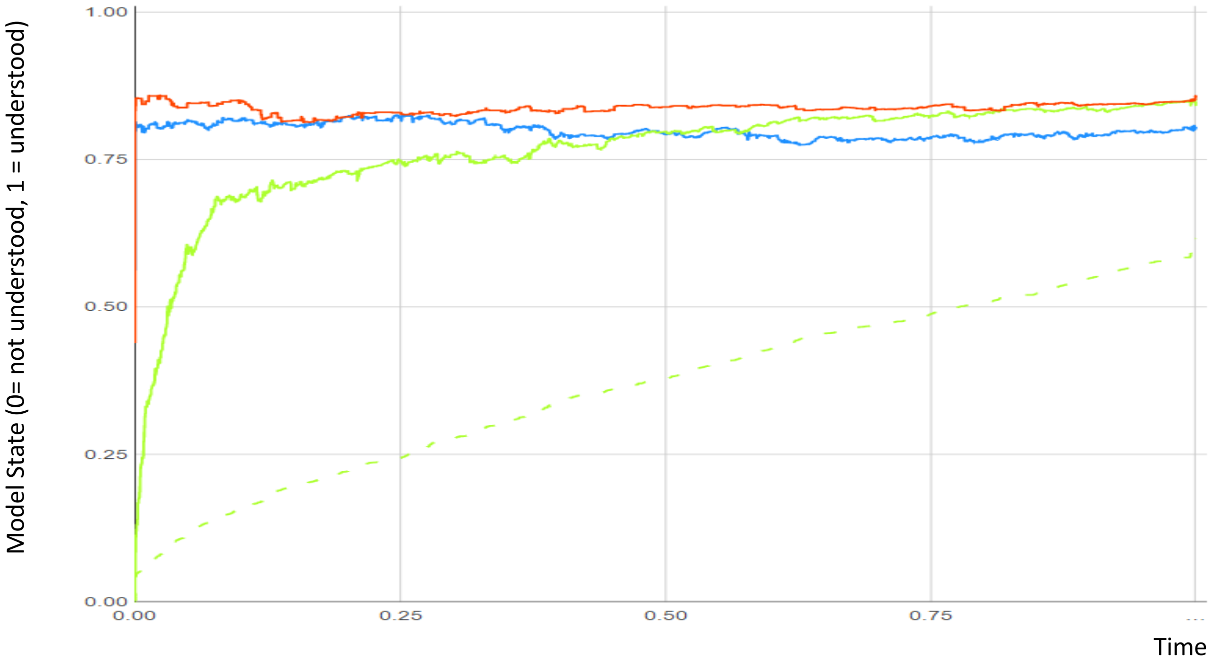
With the contextual information of the composition of the behavioural and diagnostic models, the analysis now moves forward to consider changes in their alignment throughout interaction, and how this relates to the extent of the use of the drill down functionality.

For the two uninstantiated groups, the initial measurement is taken at the first normalised quarter of interaction. For the instantiated groups the initial measurement is taken at the time at which participants first had access to the OLM. The final measurement for all groups is the time of the last interaction.

2. Points of Measurement (Study 2)

As per Study 1, Figure 36 shows an average across all participants for the average model state (blue line), the average self-assessment value (red line) and level of completeness of the model for those concepts that are active (green line), and the overall level of completeness of the model (dotted green line). Time on the x-axis is normalised between the first OLM interaction event (0.0) and the last interaction event (1.0) to make participants comparable. The average completeness of active concepts in the model starts to plateau early in the first quarter, however the overall completeness of the model continues to steadily build across the whole period of the study. A direct like-for-like comparison with equal precision is therefore not achievable with this dataset – this is a limitation. In view of the completeness information in Figure 36, the thesis considers a comparison between the end of the first quarter and the final state, for each participant. It is important to note, for the purposes of comparison, that as per Chapter 6, overall the models are not complete at any point over the studies.

Study 2. State and completeness of combined student behavioural and diagnostic models normalised over the full period of interaction.



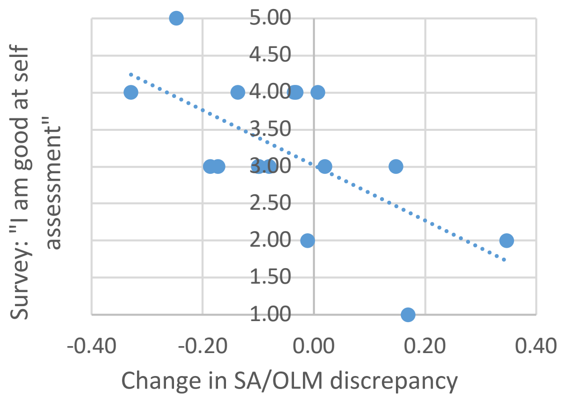
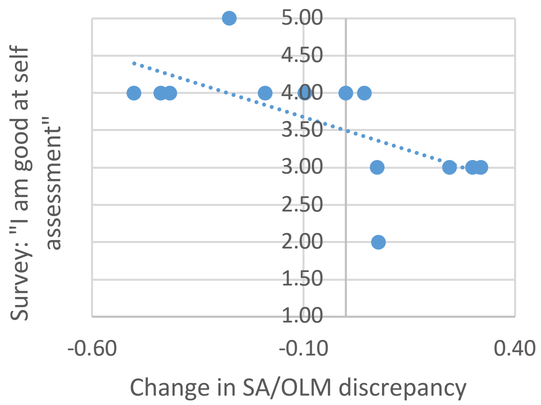
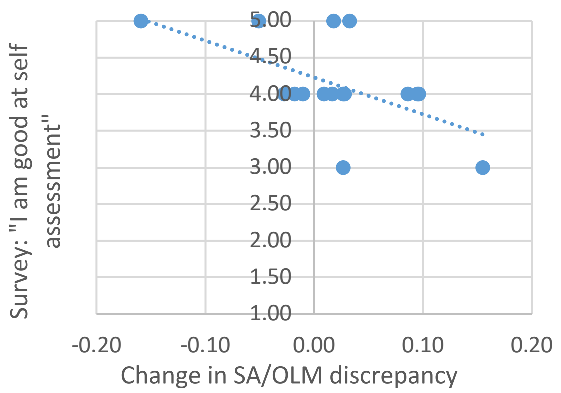
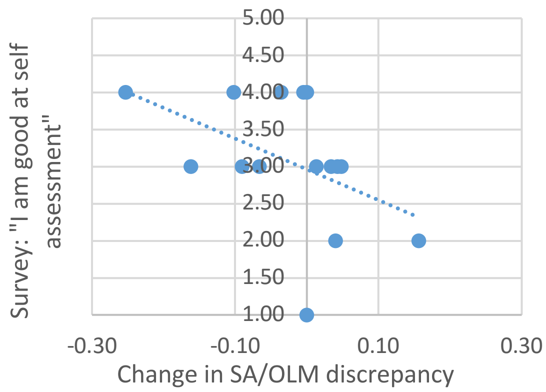
Key:
Blue – behavioural - learner model state for **knowledge** (0.0 weak, 1.0 strong)
Red – diagnostic - model of the student's self-assessed **confidence** in their ability (0.0 weak, 1.0 strong)
Solid Green – **completeness** of the model (**active concepts**) (0.0 incomplete, 1.0 fully complete)
Dotted Green – **completeness** of the model (**absolute**) (0.0 incomplete, 1.0 fully complete)

Figure 36: state and completeness of the behavioural and diagnostic models over time

APPENDIX 6: SUPPORTING DATA FOR STUDY 1

1. Q2(a) Will student self-assessment accuracy increase over time when using an open learner model with a drill down approach: changes in self-assessment accuracy Vs. student self-assessments

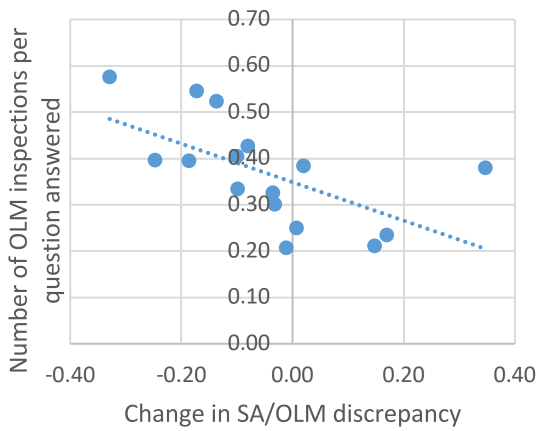
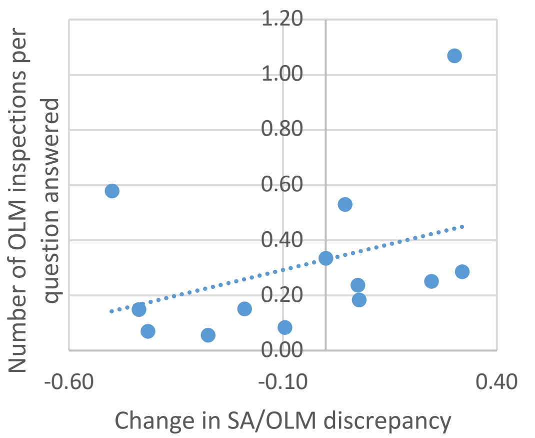
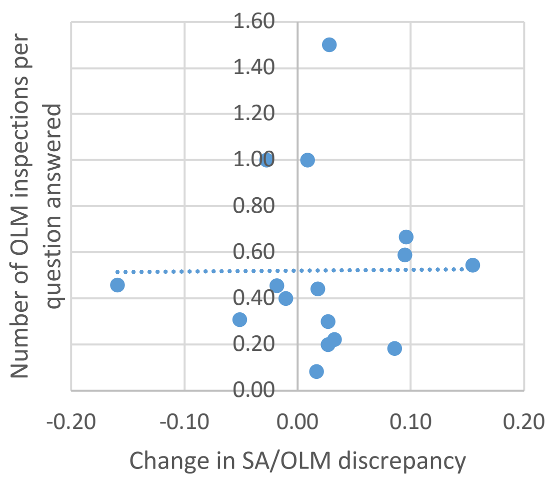
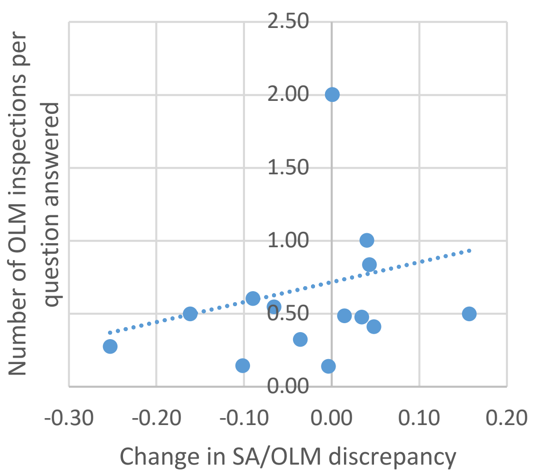
Table 38: change in behavioural/diagnostic model discrepancy, Vs self-assessment ability.

<table>
<tr><th></th><th>Mathematics</th><th>Boolean</th></tr>
<tr><td>Uninstantiated</td><td>MA(U) Group. Correlation between change in SA accuracy and students’ stated SA ability
p<0.05 significant (Spearman rank correlation)

</td><td>BO(U) Group. Correlation between SA accuracy and number of OLM inspections per question answered
p<0.01 significant (Spearman rank correlation)

</td></tr>
<tr><td>Instantiated</td><td>MA(I) Group. Correlation between SA accuracy and number of OLM inspections per question answered
p<0.05 significant (Spearman rank correlation)

</td><td>BO(I) Group. Correlation between SA accuracy and number of OLM inspections per question answered
p<0.05 significant (Spearman rank correlation)

</td></tr>
</table>

The above analysis supports the narrative included in Section 7.2.1 . It visualises the correlations between students’ ability increase in the accuracy of their self-assessments during interaction against the frequency of use of the OLM, as determined by the number of inspections per question answered. Significances refer to Spearman rank correlation.

2. Q2(a) Will student self-assessment accuracy increase over time when using an open learner model with a drill down approach: changes in self-assessment accuracy Vs. frequency of OLM use

Table 39: change in behavioural/diagnostic model discrepancy Vs inspection frequency.

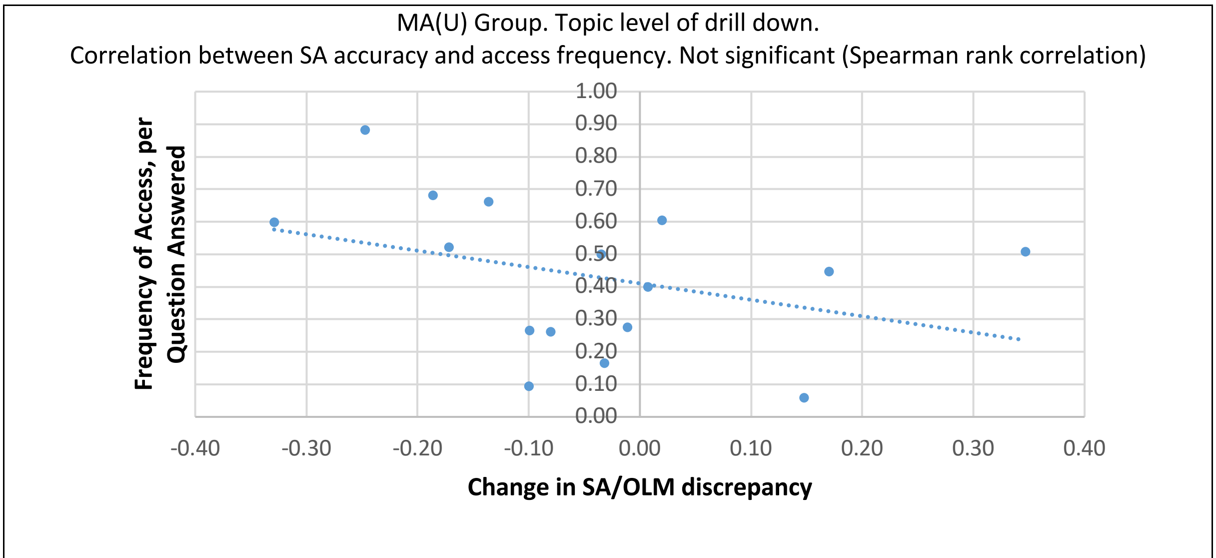
<table>
<tr><th></th><th>Mathematics</th><th>Boolean</th></tr>
<tr><td>Uninstantiated</td><td>MA(U) Group. Correlation between SA accuracy and number of OLM inspections per question answered p<0.001 significant (Spearman rank correlation)
</td><td>BO(U) Group. Correlation between SA accuracy and number of OLM inspections per question answered. Not significant (Spearman rank correlation)
</td></tr>
<tr><td>Instantiated</td><td>MA(I) Group. Correlation between SA accuracy and number of OLM inspections per question answered. Not significant (Spearman rank correlation)
</td><td>MA(U) Group. Correlation between SA accuracy and number of OLM inspections per question answered. Not significant (Spearman rank correlation)
</td></tr>
</table>

The above analysis supports the narrative included in Section 7.2.1 . It visualises the correlations between students' ability increase in the accuracy of their self-assessments during interaction against the frequency of their use of the OLM. Significances refer to Spearman rank correlation.

3. Q2(b) Will student self-assessment accuracy increase to a greater extent when greater use is made of drill down: changes in self-assessment accuracy Vs frequency of drill down access

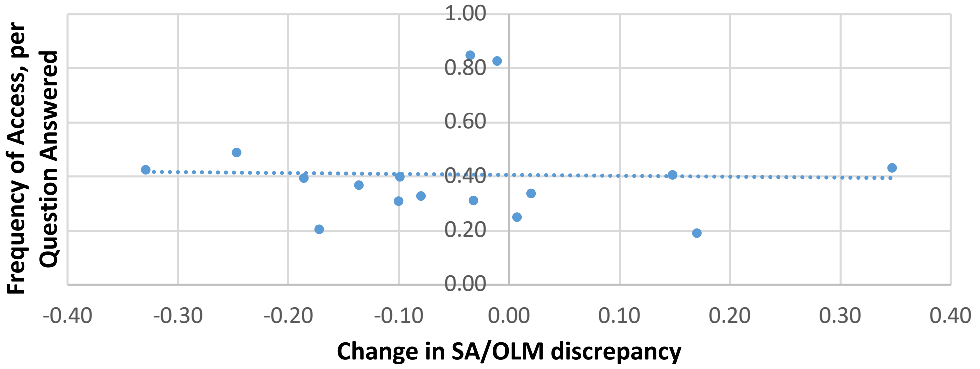
The following analysis supports the narrative included in Section 7.2.2 . It visualises the correlations between students' ability increase in the accuracy of their self-assessments during interaction against the frequency of their use of different levels of drill down. Significances refer to Spearman rank correlation.

Table 40: change in behavioural/diagnostic model discrepancy Vs OLM element inspection.



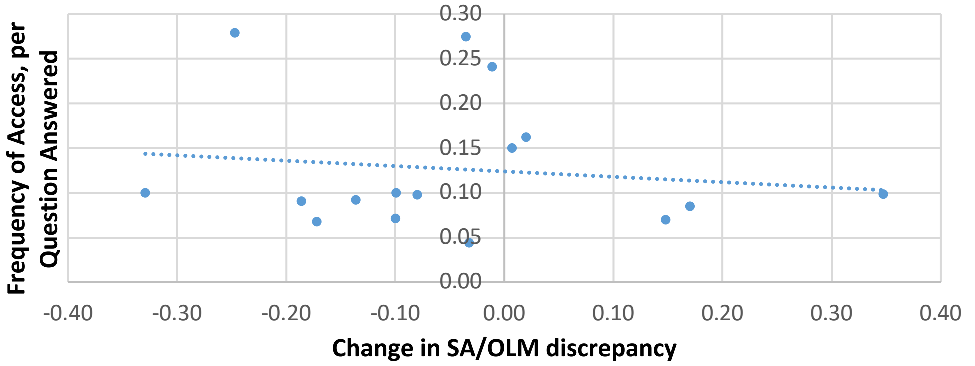
MA(U) Group. Concept level of drill down.

Correlation between SA accuracy and access frequency. Not significant (Spearman rank correlation)



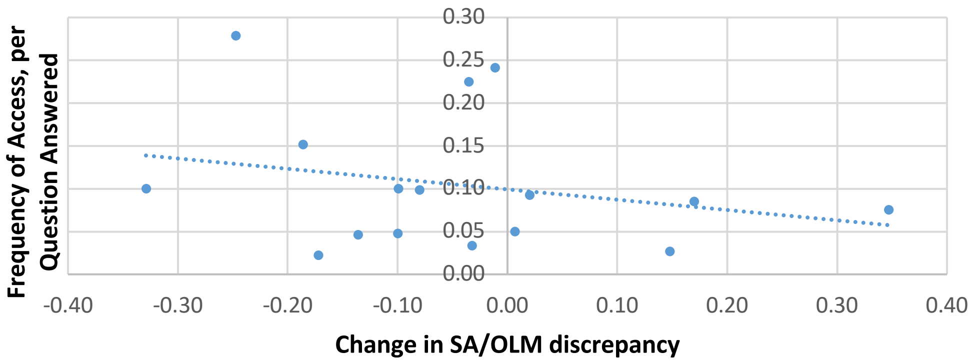
MA(U) Group. Belief level of drill down.

Correlation between SA accuracy and access frequency. Not significant (Spearman rank correlation)



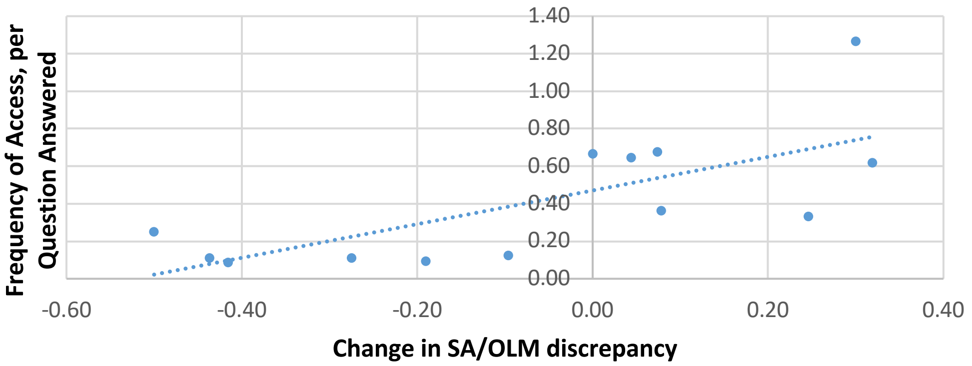
MA (U) Group. Domain level of drill down.

Correlation between SA accuracy and access frequency. Not significant (Spearman rank correlation)

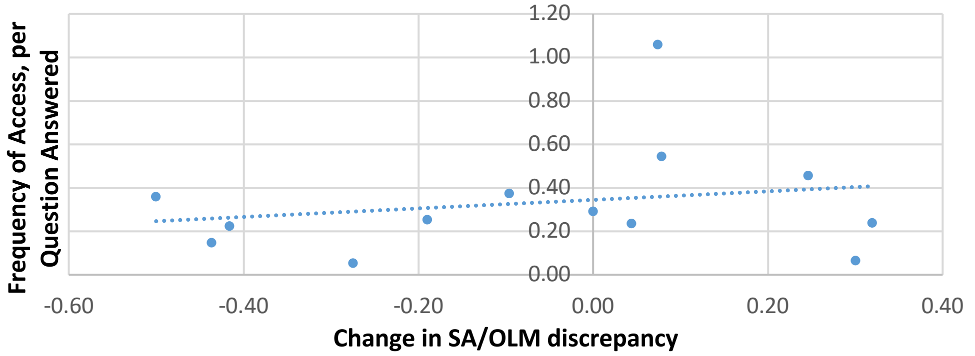


BO(U) Group. Topic level of drill down.

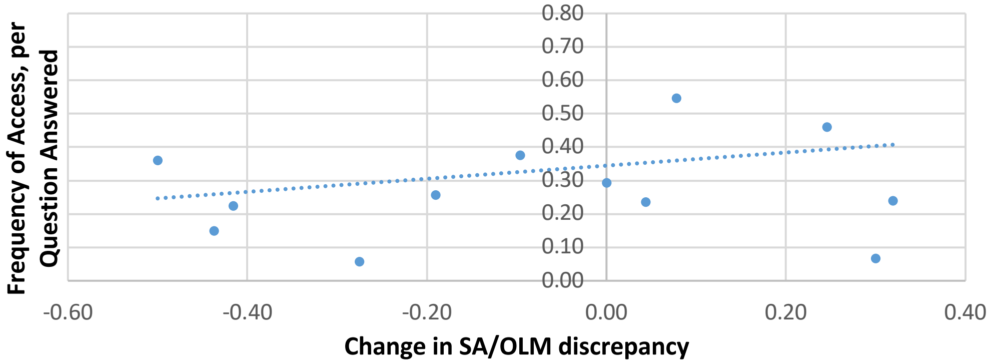
Correlation between SA accuracy and access frequency. **p<0.01 significant** (Spearman rank correlation)



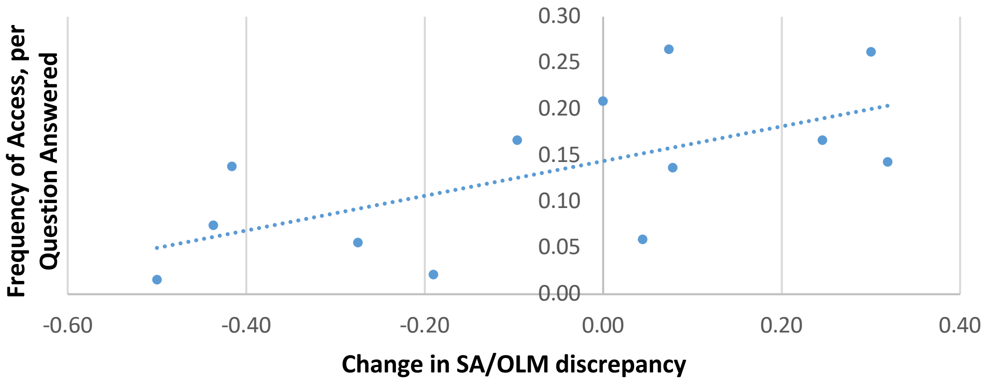
BO(U) Group. Concept level of drill down.
Correlation between SA accuracy and access frequency. Not significant (Spearman rank correlation)



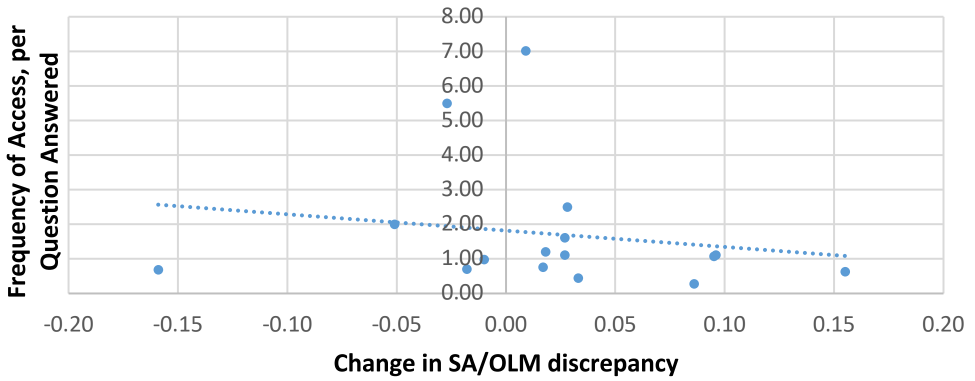
BO(U) Group. Belief level of drill down.
Correlation between SA accuracy and access frequency. **p<0.01 significant** (Spearman rank correlation)



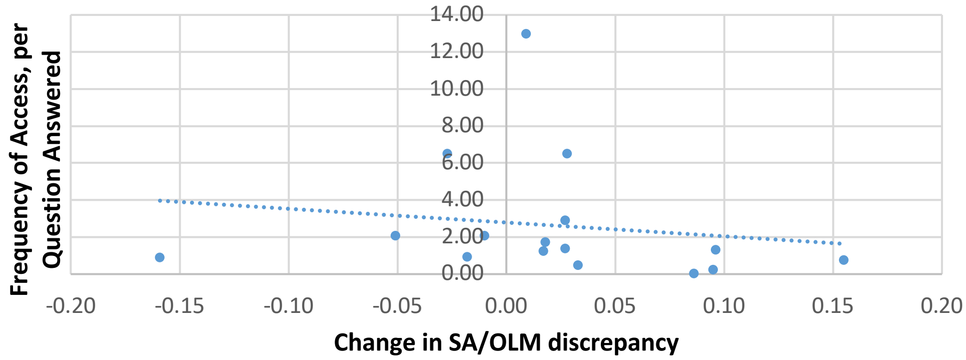
BO(U) Group. Domain level of drill down.
Correlation between SA accuracy and access frequency. **p<0.05 significant** (Spearman rank correlation)



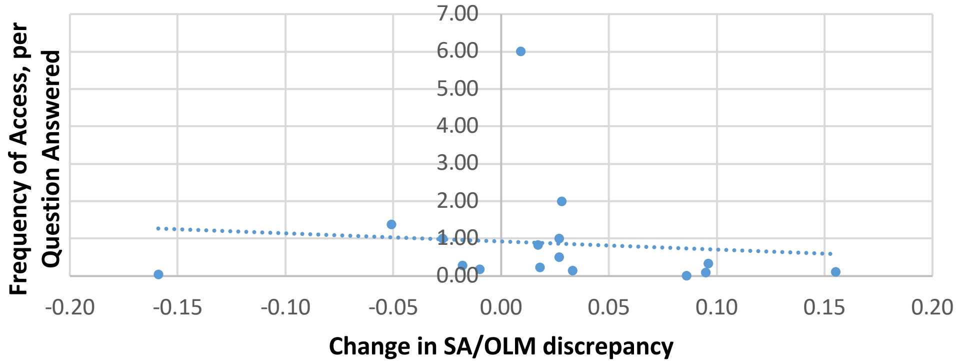
MA(I) Group. Topic level of drill down.
Correlation between SA accuracy and access frequency. Not significant (Spearman rank correlation)



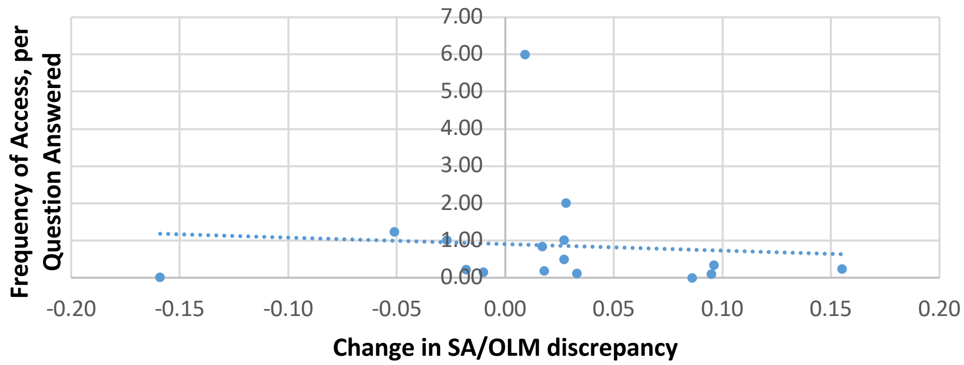
MA(I) Group. Concept level of drill down.
Correlation between SA accuracy and access frequency. Not significant (Spearman rank correlation)



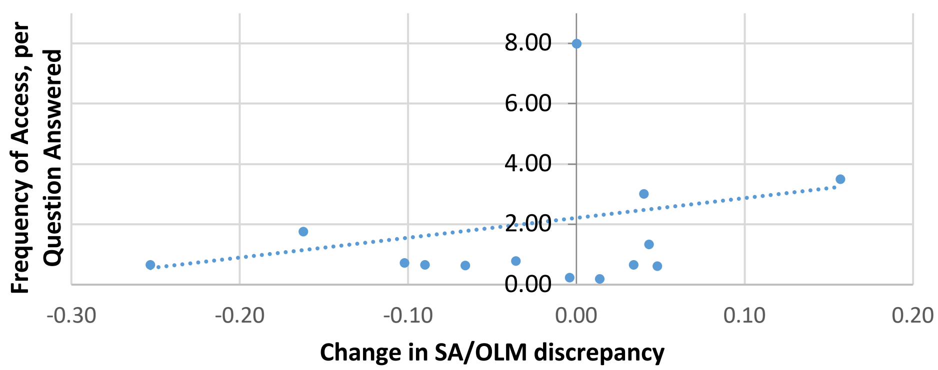
MA(I) Group. Belief level of drill down.
Correlation between SA accuracy and access frequency. Not significant (Spearman rank correlation)



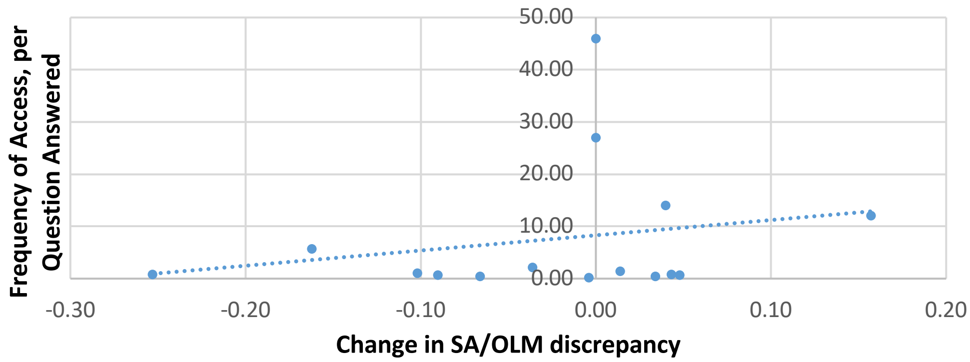
MA(I) Group. Domain level of drill down.
Correlation between SA accuracy and access frequency. Not significant (Spearman rank correlation)



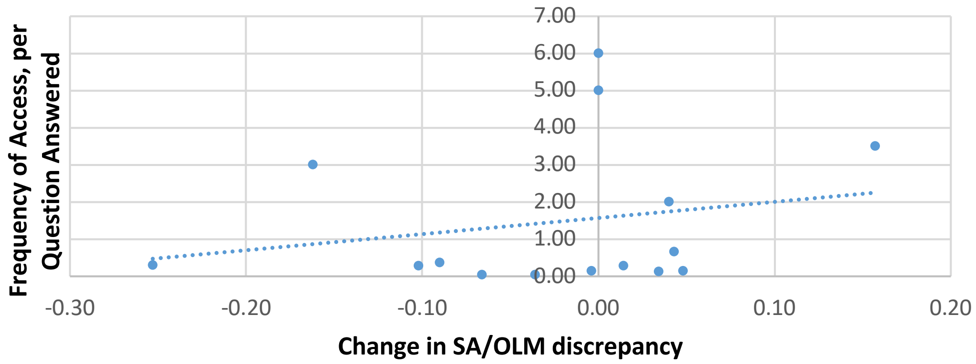
BO(I) Group. Topic level of drill down.
Correlation between SA accuracy and access frequency. Not significant (Spearman rank correlation)



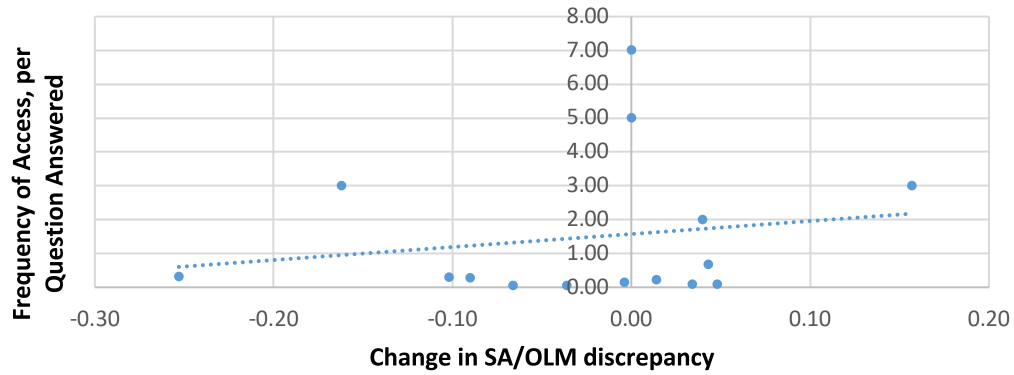
BO(I) Group. Concept level of drill down.
Correlation between SA accuracy and access frequency. Not significant (Spearman rank correlation)



BO(I) Group. Belief level of drill down.
Correlation between SA accuracy and access frequency. Not significant (Spearman rank correlation)



BO(I) Group. Domain level of drill down.
Correlation between SA accuracy and access frequency. Not significant (Spearman rank correlation)



4. Q3(a) Will students use drill down in the OLM to inspect information about problems: behavioural model state across time and OLM inspection

The following data table is used in the analysis presented in Section 7.3.1 . Students' use of drill down is more likely with items that are below the average state. Students will drill deeper with problems of greater magnitude.

Table 41: behavioural model state and deviation from the model state at point of access.

Measure	**Inspected**	**MA (U)**	**BO (U)**	**MA (I)**	**BO (I)**
Mean state of the model at the point o inspection	Overall	0.74 (SD 0.14)	0.79 (SD 0.19)	0.67 (SD 0.11)	0.67 (SD 0.26)
	Topics	0.72 (SD 0.15)	0.80 (SD 0.18)	0.67 (SD 0.10)	0.67 (SD 0.26)
	Concepts	0.74 (SD 0.15)	0.80 (SD 0.17)	0.69 (SD 0.12)	0.67 (SD 0.26)
	Beliefs	0.76 (SD 0.15)	0.76 (SD 0.18)	0.67 (SD 0.11)	0.65 (SD 0.27)
	Domain	0.77 (SD 0.15)	0.76 (SD 0.19)	0.66 (SD 0.11)	0.64 (SD 0.27)
Mean state of mode element inspected	Topics	0.71 (SD 0.14)	0.84 (SD 0.19)	0.72 (SD 0.15)	0.67 (SD 0.25)
	Concepts	0.62 (SD 0.21)	0.69 (SD 0.23)	0.69 (SD 0.18)	0.70 (SD 0.19)
	Beliefs	0.48 (SD 0.20)	0.55 (SD 0.35)	0.54 (SD 0.22)	0.60 (SD 0.24)
	Domain	0.51 (SD 0.29)	0.50 (SD 0.36)	0.57 (SD 0.20)	0.57 (SD 0.27)
Mean difference between inspected state and average state	Topics	-0.02 (SD 0.05)	0.04 (SD 0.11)	0.04 (SD 0.07)	0.00 (SD 0.04)
	Concepts	-0.12 (SD 0.17)	-0.11 (SD 0.14)	0.00 (SD 0.10)	0.03 (SD 0.14)
	Beliefs	-0.27 (SD 0.23)	-0.21 (SD 0.40)	-0.13 (SD 0.20)	-0.05 (SD 0.29)
	Domain	-0.25 (SD 0.31)	-0.26 (SD 0.40)	-0.10 (SD 0.20)	-0.07 (SD 0.29)
Significance	Topics*	Not significant	Not significant	Not significant	Not significant
	Concepts*	Not significant	Not significant	Not significant	Not significant
	Beliefs*	**p<0.05**	Not significant	Not significant	Not significant
	Domain*	Not significant	Not significant	Not significant	Not significant
	Between Levels of Drill Down**	**p<0.01**	**p<0.05**	**p<0.01**	Not significant

*significances between mean state and state inspected calculated with a two-tailed T-test.
**significances between different drill down levels calculated with a 1 way ANOVA (related)

5. Q3(b) Will students use drill down in the OLM to inspect information about areas of uncertainty: diagnostic model state across time and OLM inspection

The following data table is used in the analysis presented in Section 7.3.2 . Students' use of drill down is not statistically more likely with items where confidence is below that of the average model state. Students' use of drill down is not related to their underlying confidence of the measure.

Table 42: diagnostic model state and deviation from the model state at the point of access.

Measure	Inspected	MA (U)	BO (U)	MA (I)	BO (I)
Mean state of the confidence model at the point of inspection	Overall	0.52 (SD 0.18)	0.66 (SD 0.29)	0.69 (SD 0.15)	0.63 (SD 0.23)
	Topics	0.51 (SD 0.18)	0.70 (SD 0.29)	0.68 (SD 0.15)	0.63 (SD 0.22)
	Concepts	0.53 (SD 0.18)	0.70 (SD 0.30)	0.70 (SD 0.15)	0.63 (SD 0.22)
	Beliefs	0.53 (SD 0.18)	0.74 (SD 0.30)	0.69 (SD 0.13)	0.61 (SD 0.20)
	Domain	0.52 (SD 0.20)	0.73 (SD 0.31)	0.68 (SD 0.14)	0.60 (SD 0.20)
Mean state of confidence model for element inspected	Topics	0.53 (SD 0.19)	0.72 (SD 0.30)	0.73 (SD 0.26)	0.56 (SD 0.27)
	Concepts	0.58 (SD 0.24)	0.72 (SD 0.26)	0.75 (SD 0.26)	0.64 (SD 0.23)
	Beliefs	0.57 (SD 0.20)	0.65 (SD 0.33)	0.72 (SD 0.24)	0.60 (SD 0.21)
	Domain	0.65 (SD 0.25)	0.61 (SD 0.35)	0.72 (SD 0.27)	0.59 (SD 0.22)
Mean difference betwe inspected confidence state and average confidence state	Topics	0.02 (SD 0.07)	0.02 (SD 0.08)	0.05 (SD 0.13)	-0.03 (SD 0.14)
	Concepts	0.05 (SD 0.13)	0.03 (SD 0.15)	0.05 (SD 0.15)	0.01 (SD 0.17)
	Beliefs	0.04 (SD 0.14)	-0.09 (SD 0.31)	0.04 (SD 0.16)	-0.01 (SD 0.21)
	Domain	0.13 (SD 0.24)	-0.12 (SD 0.35)	0.04 (SD 0.17)	-0.01 (SD 0.20)
Significance	Topics*	Not significant	Not significant	Not significant	Not significant
	Concepts*	Not significant	Not significant	Not significant	Not significant
	Beliefs*	Not significant	Not significant	Not significant	Not significant
	Domain*	Not significant	Not significant	Not significant	Not significant
	Between Levels of Drill Down**	Not significant	Not significant	Not significant	Not significant

*significances between mean state and state inspected calculated with a two-tailed T-test.
**significances between different drill down levels calculated with a 1 way ANOVA (related)

6. Q3(c) Will students use an OLM with a drill down approach to focus on one domain area at a time?

The following data supplements the analysis summarised in Section 7.3.3 .

Table 43: how many question blocks ago was the same concept worked with.

Measure	Inspected	MA (U)	BO (U)	MA (I)	BO (I)
Mean number of question blocks ago the same topic/concept encountered.	Topics	3.72 (SD 1.71)	1.98 (SD 1.17)	2.00 (SD 1.14)	1.70 (SD 0.72)
	Concepts	2.52 (SD 1.47)	2.01 (SD 0.75)	2.79 (SD 1.89)	2.54 (SD 1.68)
	Beliefs	1.80 (SD 1.14)	1.48 (SD 0.42)	2.08 (SD 1.42)	1.64 (SD 0.61)
	Domain	1.66 (SD 1.11)	1.69 (SD 0.66)	2.22 (SD 2.09)	1.78 (SD 0.77)
Significance between levels of drill down*		**p<0.001**	Not significant	Not significant	Not significant

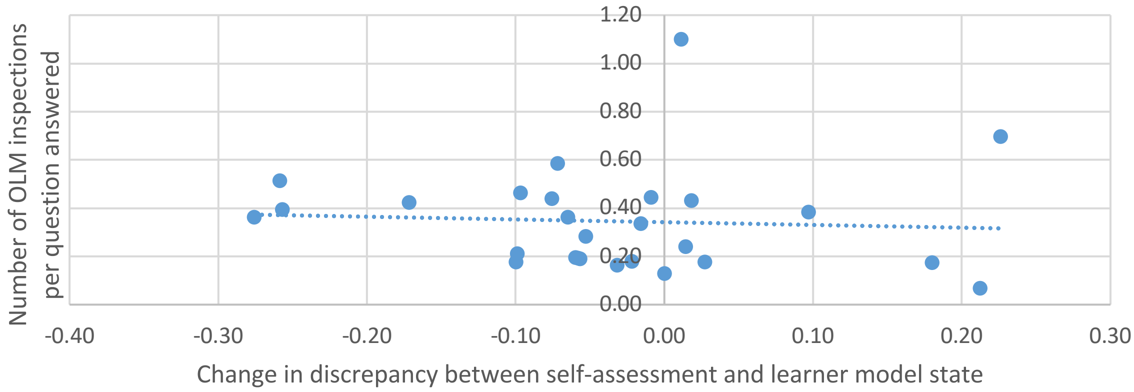
*significances between different drill down levels calculated with a 1 way ANOVA (related)

APPENDIX 7: SUPPORTING DATA FOR STUDY 2

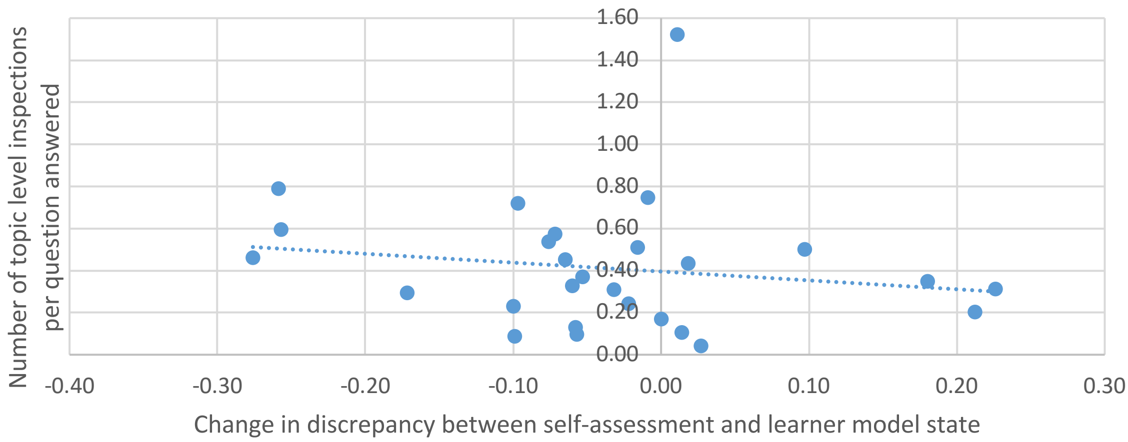
1. Q2(a) Will student self-assessment accuracy increase over time when using an open learner model with a drill down approach: changes in self-assessment accuracy Vs. OLM inspection frequency

The following analysis supports the narrative included in Section 8.2.1 . It visualises the correlations between students' ability increase in the accuracy of their self-assessments during interaction against the frequency of use of the OLM, as determined by the number of inspections per question answered. Significances refer to Spearman rank correlation.

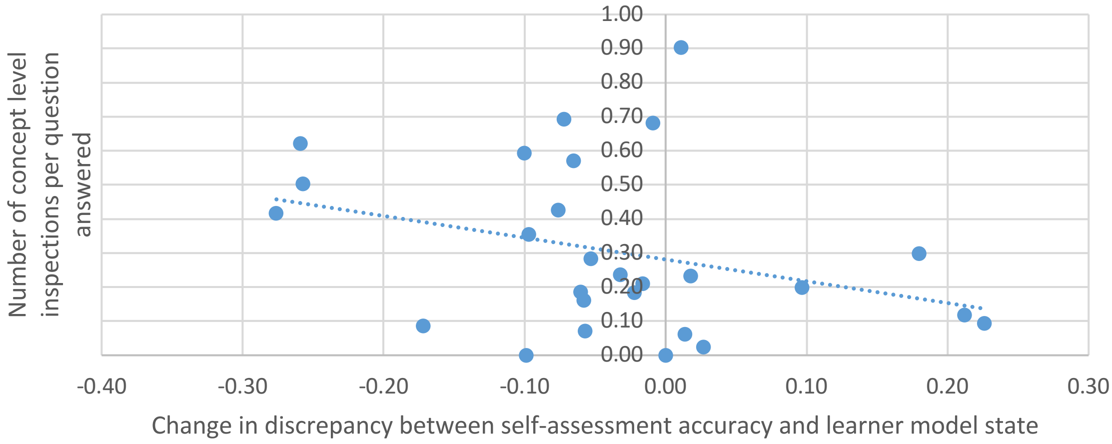
(a) Overall. Not significant*



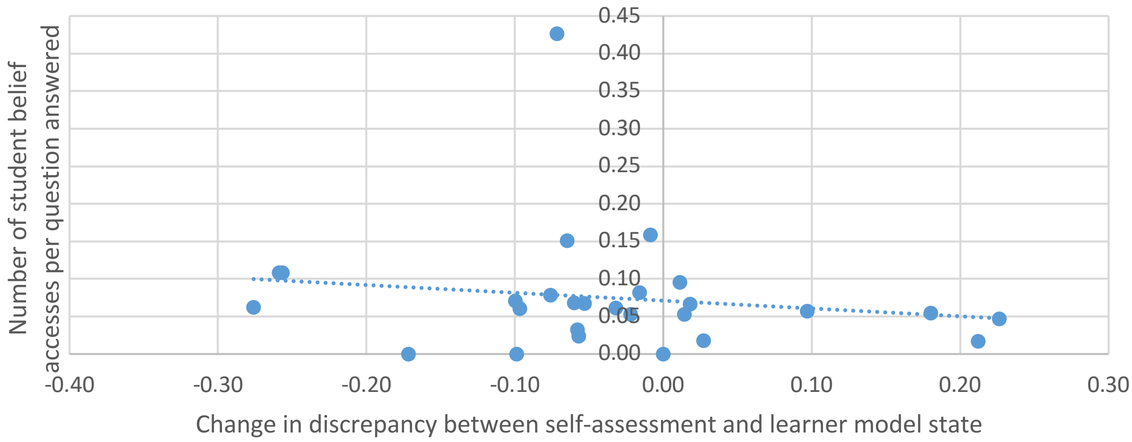
(b) Topic level inspections. Not significant*



(c) Concept level inspections. Not significant*



(d) Belief level inspections. Not significant*



(e) Domain level inspections. Not significant*

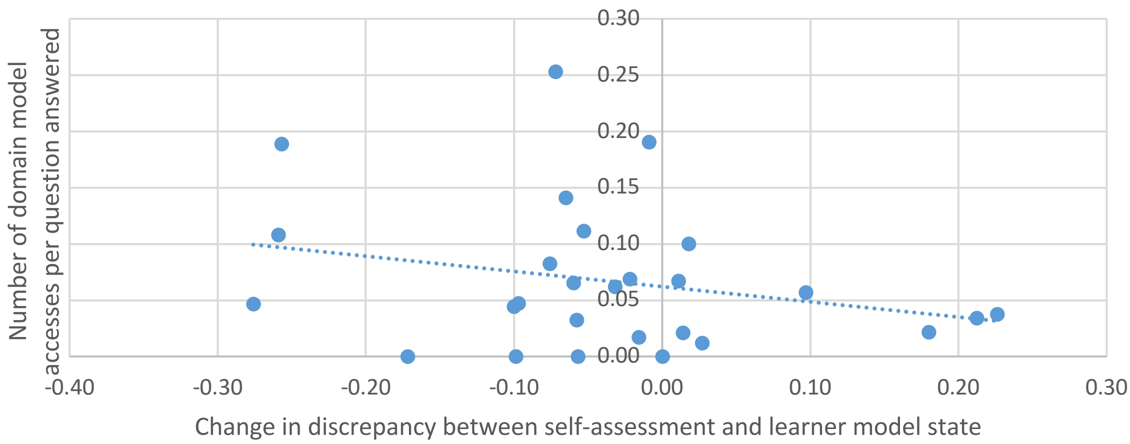


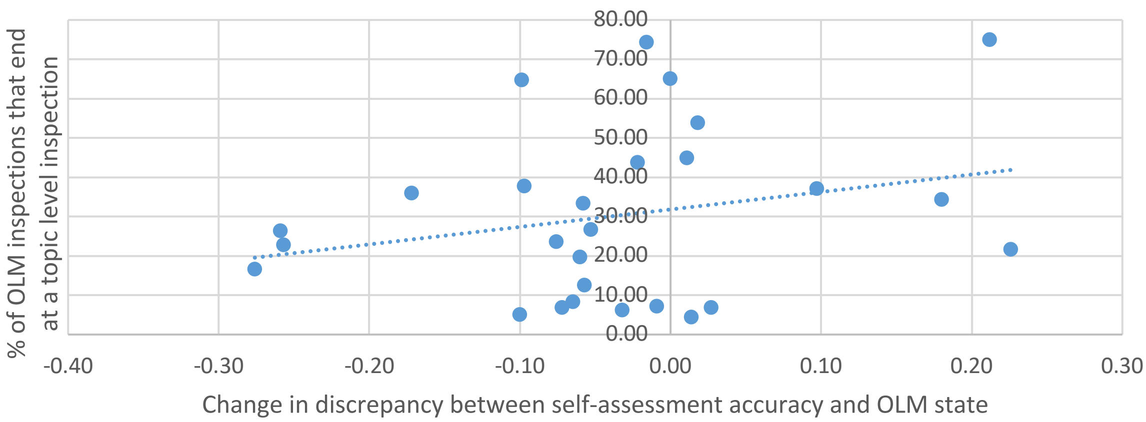
Figure 37: changes in self-assessment discrepancy, Vs frequency of inspection of model.
*Significances refer to Spearman rank correlation.

2. Q2(b) Will student self-assessment accuracy increase to a greater extent when greater use is made of drill down: changes in self-assessment accuracy Vs. use of drill down

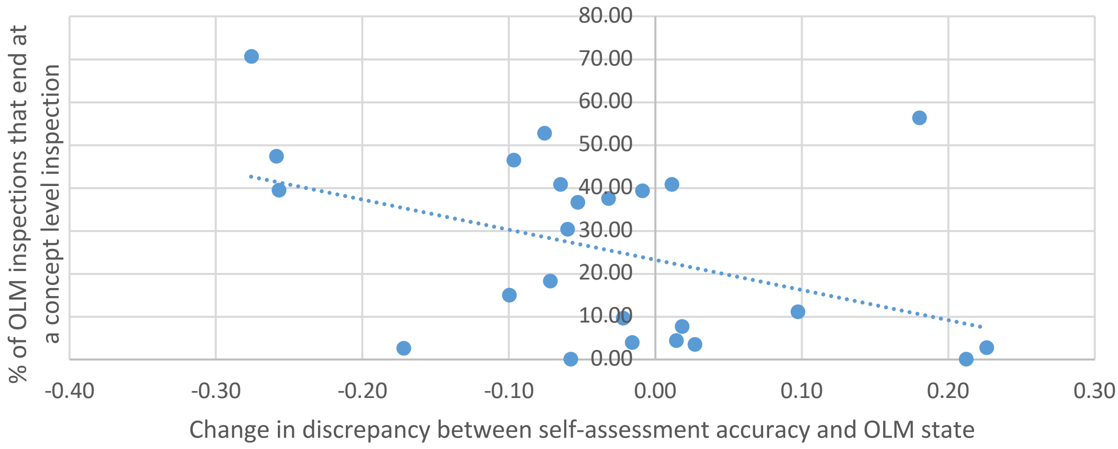
The following analysis supports the narrative included in Section 8.2.2 . It visualises the correlations between students' ability increase in the accuracy of their self-assessments during interaction against the proportion of their inspections of the learner model were to a given depth of drill down. Significances refer to Spearman rank correlation.

(a) Overall. Not significant*

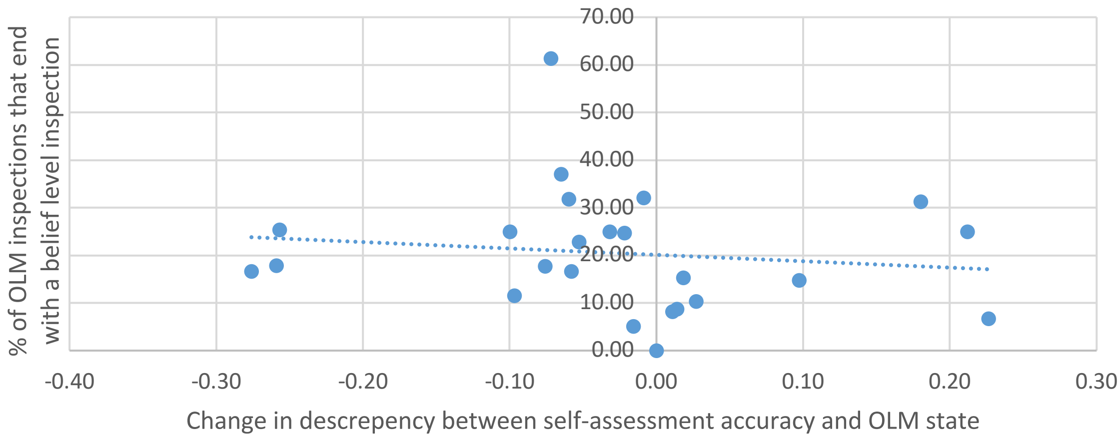
(b) Topic level inspections. Not significant*



(c) Concept level inspections. **p<0.05 (significant)***



(d) Belief level inspections. **p<0.05 (significant)***



(e) Domain level inspections. Not significant*

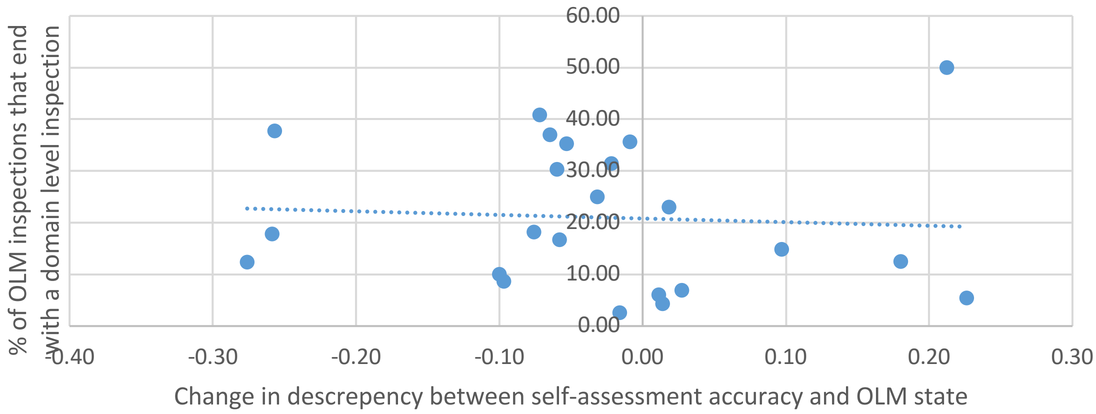


Figure 38: changes in self-assessment discrepancy, compared to depth of inspection.
*Significances refer to Spearman rank correlation.

3. Q3(a) Will students use drill down in the OLM to inspect information about problems: behavioural model state across time and OLM inspection

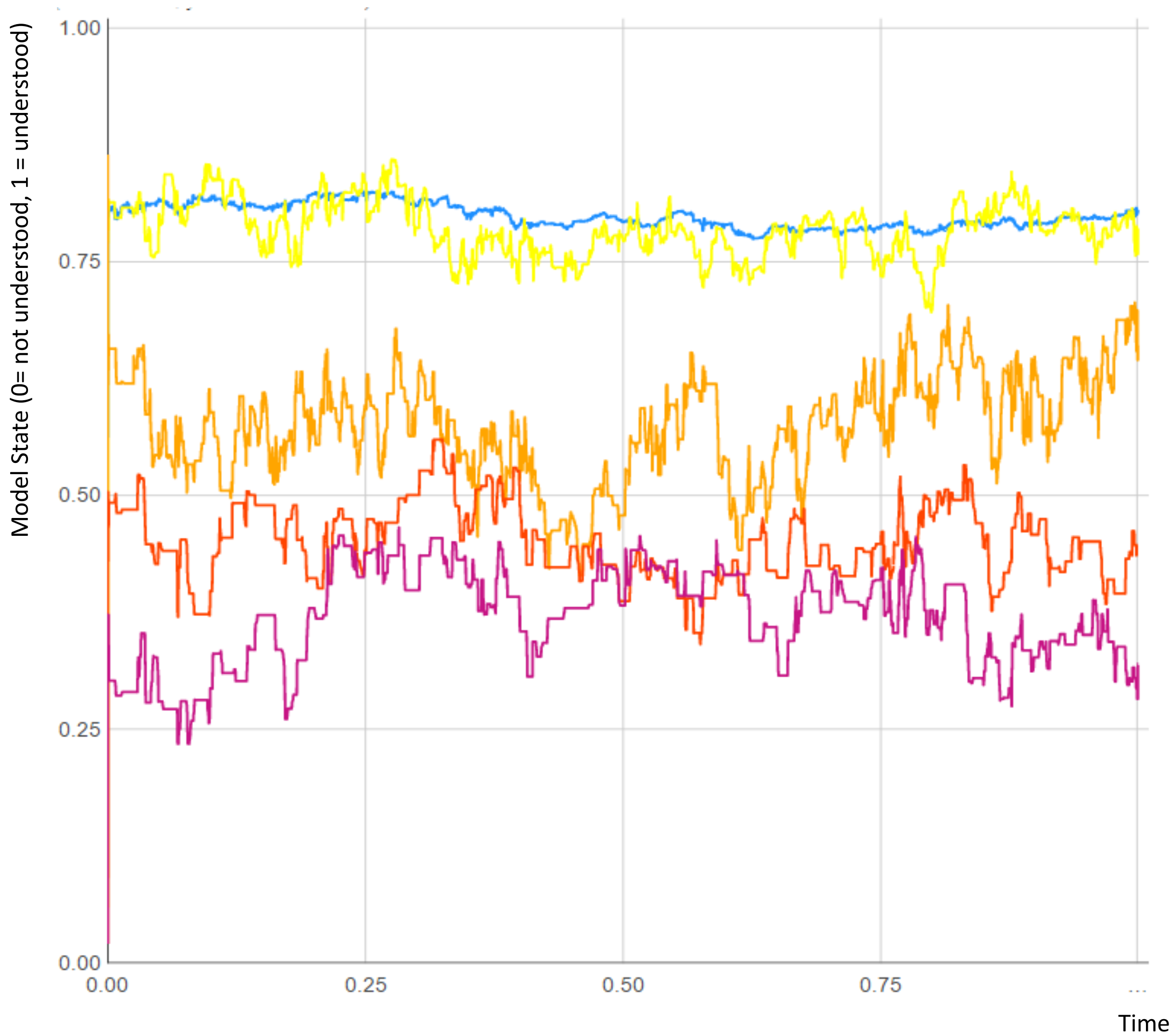
The following data table is used in the analysis presented in Section 8.3.1 . Students' use of drill down is more likely with items that are below the average state. Students will also drill deeper with problems of greater magnitude.

Table 44: behavioural model state and deviation from the model state at point of access.

Measure	At the Point of Inspection					Significance**
	Overall	Topics	Concepts	Beliefs	Domain	
Total inspections	2023	2606	1991	440	465	
Inspections per participant	74.92 (SD 75.52)	96.52 (SD 103.18)	73.74 (SD 86.52)	16.30 (SD 17.48)	17.22 (SD 25.09)	
Mean state of the overall model	0.80 (SD 0.14)	0.80 (SD 0.13)	0.81 (SD 0.12)	0.81 (SD 0.11)	0.80 (SD 0.12)	Not significant
Mean state of the model element inspected	n/a whole model on view	0.77 (SD 0.15)	0.59 (SD 0.23)	0.49 (SD 0.25)	0.43 (0.25)	**p<0.000001**
Mean difference from state of the element and state of the overall model		-0.03 (SD 0.07)	-0.23 (SD 0.20)	-0.32 (SD 0.22)	-0.37 (SD 0.25)	**p<0.0000001**
Significance*		Not significant	**p<0.001**	**p<0.0001**	**p<0.0001**	

* Significance between model state and inspection state calculated with a two-tailed T-test.
** Significance between drill down levels calculated with a 1 way ANOVA (related).
*** SD values refer to variance between participants.

The relationship of items selected for more detailed inspection being weaker than the model is also consistent across the full period of the study. The averages are shown in Figure 39. Each learner model interaction period is normalised from 0 (start) to 1 (end). Periods of inactivity greater than 30 minutes are removed from the calculation. Inspection of topics (yellow) is close to the average, topics (orange) is below, concepts (red) below that, and domain (purple) as considerably below the average state of the model (p<0.0000001). It is important to note that topics will be a partial model (average of several concepts), and so may represent a more average state. The values for the other three metrics are to the same precision and may be directly compared.



Key:
Blue – average model state at the point of inspection.
Yellow – inspection of the model – topic level
Orange – inspection of the model – concept level
Red – inspection of the model – belief level
Purple – inspection of domain model content.

Figure 39: behavioural model state and deviation at point of access, across time.[26]

4. Q3(b) Will students use drill down in the OLM to inspect information about areas of uncertainty: diagnostic model state across time and OLM inspection

The following data table is used in the analysis presented in Section 8.3.2 . Students' use of drill down is not statistically more likely with items where confidence is below that of the

[26] Screen shot taken from the analysis suite developed for the thesis to visualise interaction log data.

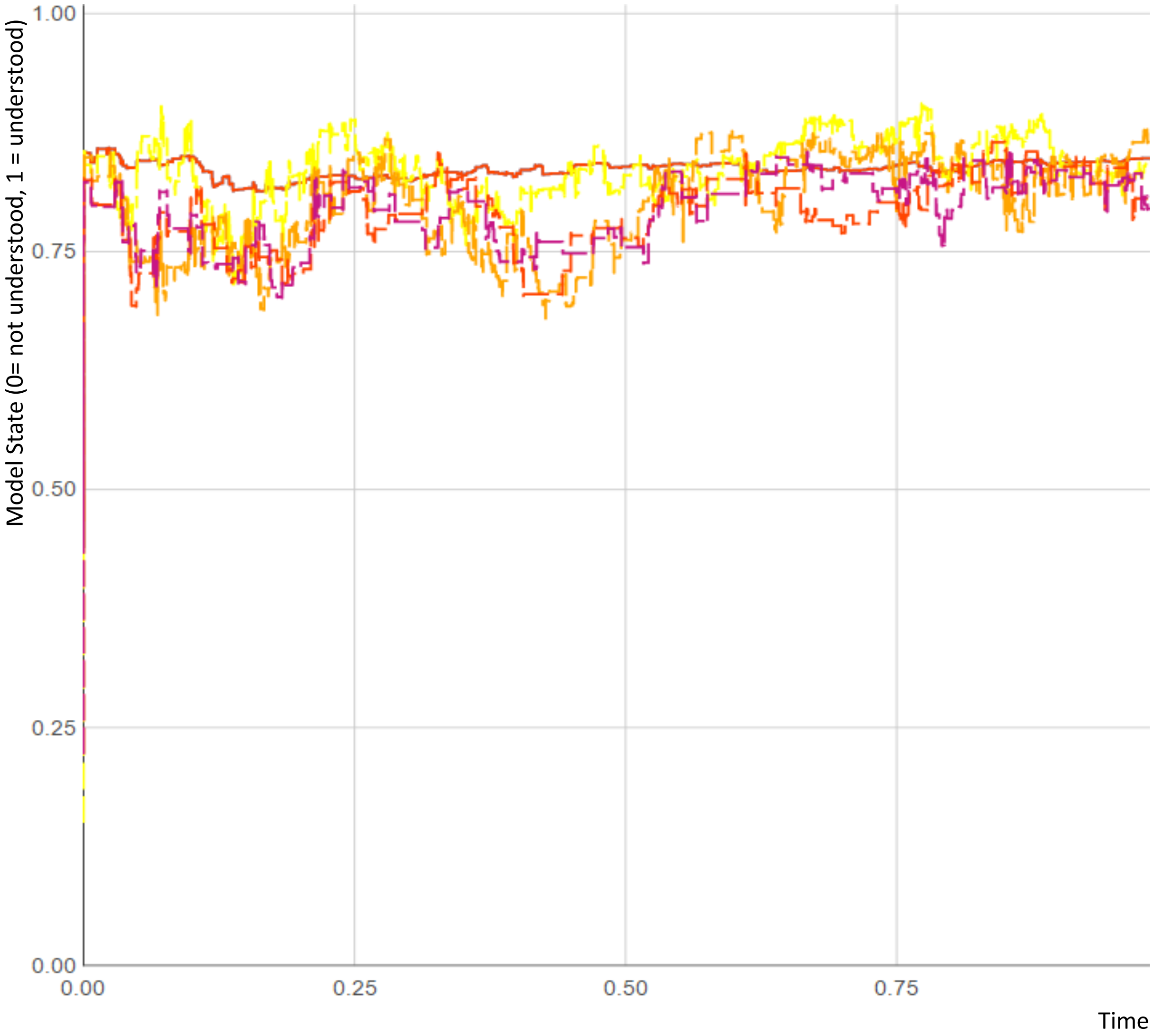
average model state. Students' use of drill down is not related to their underlying confidence of the measure.

Table 45: diagnostic model state and deviation from the model state at point of access.

Measure	At the Point of Inspection				
	Overall	Topics	Concepts	Beliefs	Domain
Total inspections	2023	2606	1991	440	465
Mean inspections per participant	74.92 (SD 75.52)	96.52 (SD 103.18)	73.74 (SD 86.52)	16.30 (SD 17.48)	17.22 (SD 25.09)
Mean state of the confidence model	0.84 (SD 0.15)	0.84 (SD 0.14)	0.83 (SD 0.14)	0.83 (SD 0.15)	0.83 (SD 0.15)
Mean state of the confidence model for items inspected	n/a whole model on view	0.83 (SD 0.16)	0.81 (SD 0.19)	0.80 (SD 0.20)	0.79 (SD 0.19)
Mean magnitude difference from mean confidence model and item inspected		-0.01 (SD 0.06)	-0.03 (SD 0.12)	-0.03 (SD 0.14)	-0.04 (SD 0.14)
Overall significance*		Not significant	Not significant	Not significant	Not significant

*Significance between model state and inspection state calculated with a two-tailed T-test.
** SD values refer to variance between participants.

The relationship of items selected for more detailed inspection showing no correlation to diagnostic model is also consistent across the full period of the study. The averages are shown in Figure 39. The solid red line is the state of the model for students' self-assessments of their ability to answer questions correctly. Throughout interaction the values for items drilled down into (yellow – topics; orange – concepts; dotted red – beliefs; dotted purple – domain) are in each case about the average state and remain consistent across time. The fact that the lines are overlaid for the most part is a visual indication of the lack of variance between the different depths of drill down.



Key:
Solid Red – average model state at the point of inspection.
Yellow – inspection of the model – topic level
Orange – inspection of the model – concept level
Dotted Red – inspection of the model – belief level
Purple – inspection of domain model content.

Figure 40: diagnostic model state and deviation at point of access, across time.[27]

[27] Screen shot taken from the analysis suite developed for the thesis to visualise interaction log data.

5. Q3(c) Will students use the OLM with a drill down approach to focus on one domain area at a time: OLM inspection Vs domain area

The supporting evidence for Section 8.3.3 is included here. It considers whether students' use of drill down is with items that have just been updated, or whether students are inspecting the formative assessment information as delayed feedback. In the majority of information students inspect in the model is from a question just answered, but there are no significant differences across levels of drill down, which suggests that the decision to drill into the model is not affected by how recently the information is updated (Table 46).

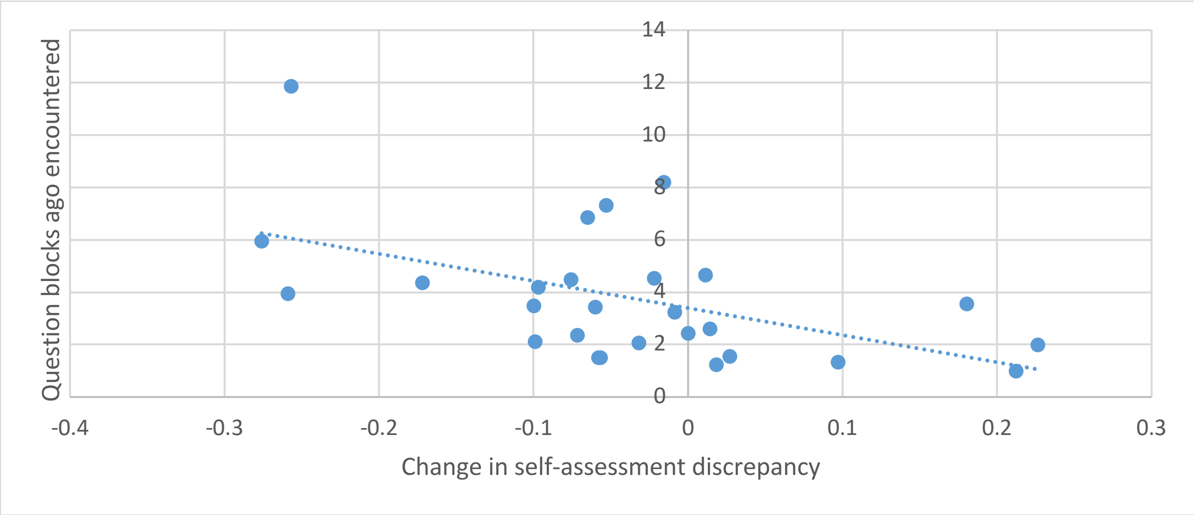
Table 46: how many question blocks ago was the same concept worked with.

	Topics	Concepts	Beliefs	Domain	Significance*
Total number of inspections	2606	1991	440	465	
Mean, per participant	3.94 (SD 2.54)	5.62 (SD 4.47)	4.40 (SD 6.16)	4.56 (SD 5.06)	Not significant
Mean Median, per participant	1.04 (SD 0.19)	1.64 (SD 1.46)	1.21 (SD 0.41)	1.22 (SD 0.75)	Not significant
% of inspections that are of concepts encountered during the **last** set of questions	68.9 (SD 14.7)	56.8 (SD 29.0)	72.4 (SD 31.6)	75.3 (SD 29.0)	Not significant

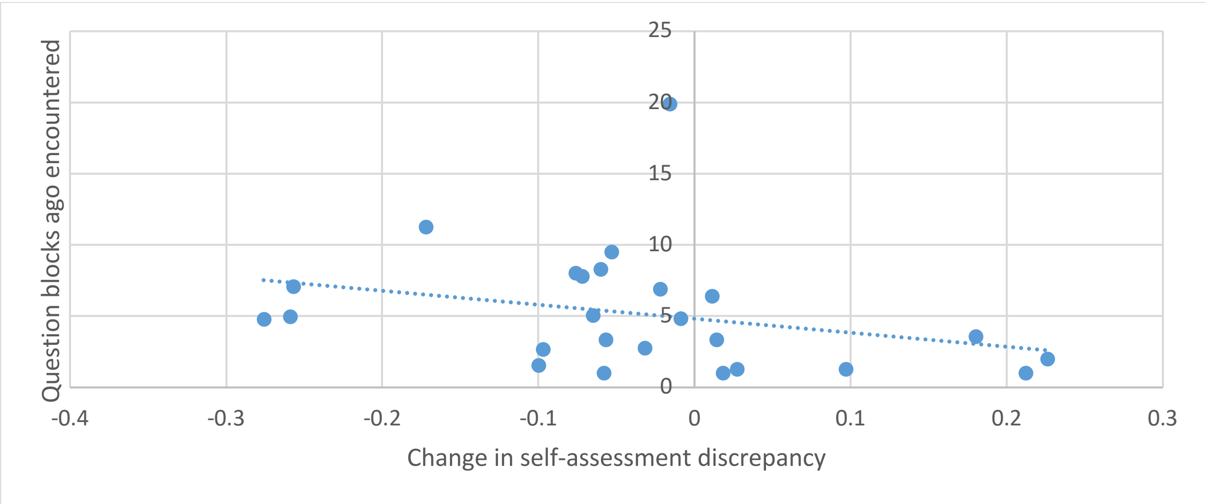
*Significance between drill down levels calculated with a 1 way ANOVA (related).

There is evidence however to suggest that those who more often inspected items that had been updated some time previously increased in self-assessment accuracy to a greater extent. Figure 41 compares the informational focus of content inspected against how recently an update attempt was made (i.e. through question attempts) and student ability to increase in self-assessment accuracy during interaction with the intervention. There is a significant relationship, at all levels of information granularity, to suggest that greater increases in self-assessment accuracy are associated with inspecting items less recently updated.

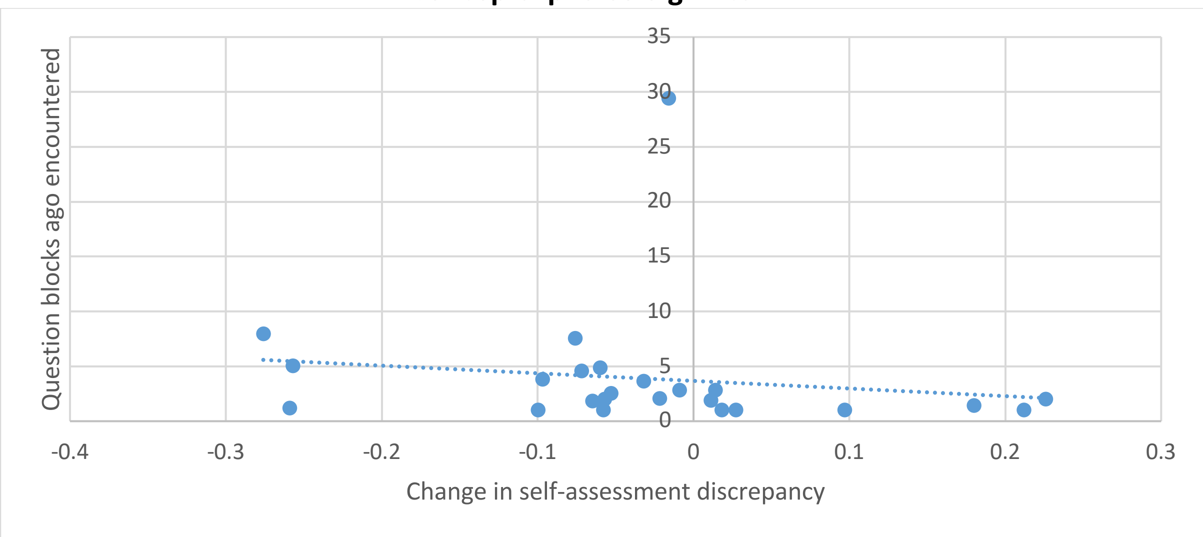
Overall. p<0.01 Significant*



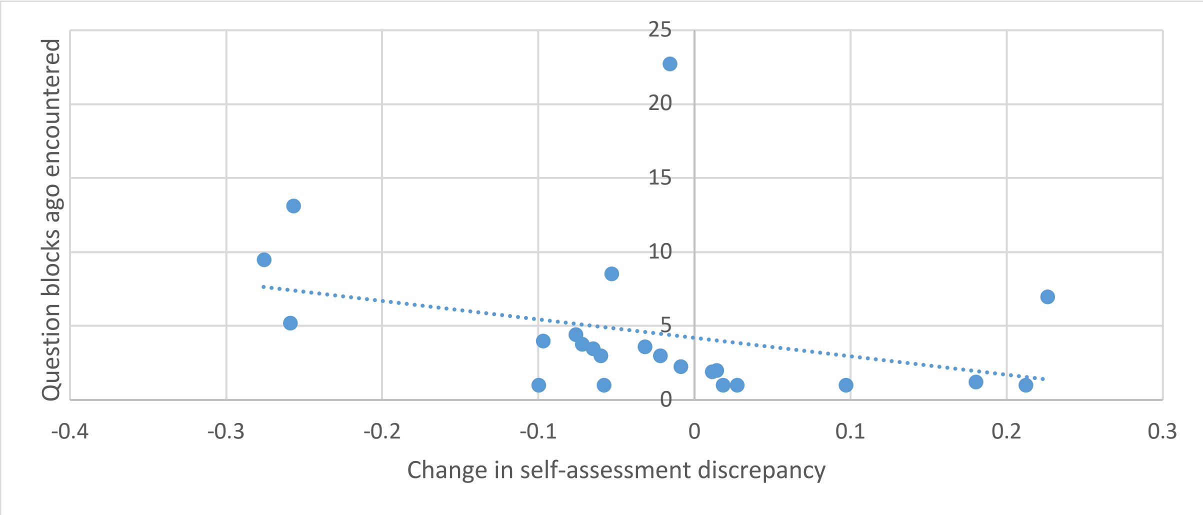
Topics. p<0.05 Significant*



Concepts. p<0.05 Significant*



Beliefs. p<0.01 Significant*



Domain. p<0.05 Significant*

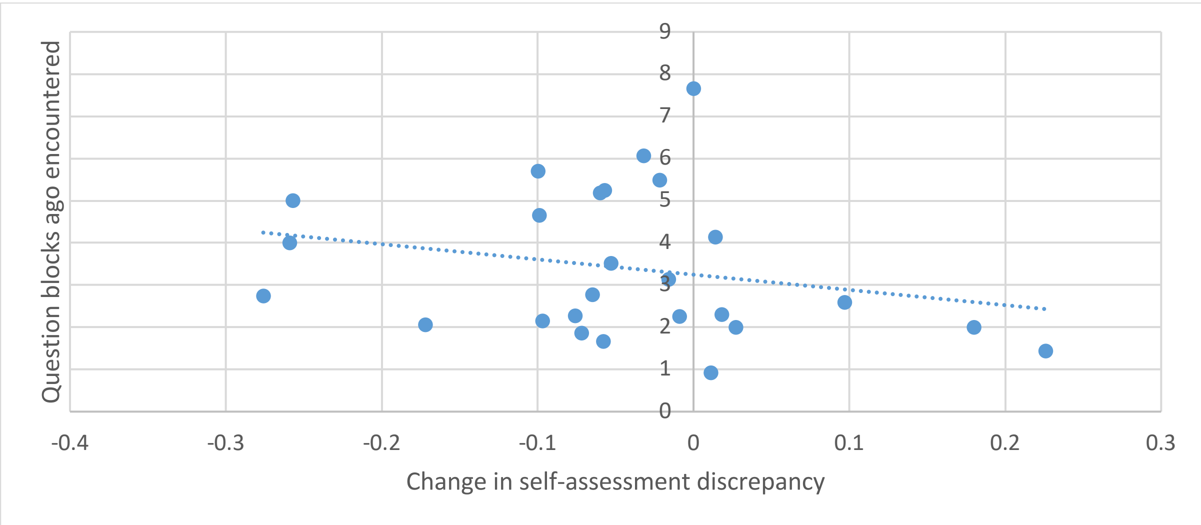


Figure 41: how recently domain content is encountered in questions, Vs change in self-assessment discrepancy.

*Significances refer to Spearman rank correlation.