Uncertainty in Open Learner Models: Visualising Inconsistencies in the Underlying Data

Lamiya Al-Shanfari
University of Birmingham, UK
lsa339@bham.ac.uk

Carrie Demmans Epp
University of Pittsburgh, USA
cdemmans@pitt.edu

Susan Bull
University College London, UK
s.bull@ucl.ac.uk

ABSTRACT

This paper suggests different methods for visualising uncertainty in open learner models (OLM). In order to visualise the uncertainty in OLMs, two factors need to be measured, namely the source of the uncertainty in the data and the level of uncertainty in the learner model. This paper proposes a method to detect the source of uncertainty within a learner model: outlier analysis is employed to identify inconsistencies in the data set from which the OLM is built. The level of uncertainty that is present in the model is determined by summing the influence weights of the learner model data that was identified as being inconsistent. Different approaches to visualising this uncertainty within OLMs are proposed; and benefits for OLMs that visualise uncertainty in learner models that can be jointly maintained by student and system, are argued.

Keywords

Uncertainty, open learner models, visualisation

1. INTRODUCTION

Learner models represent what a teaching system believes about the learner's knowledge, beliefs, competencies, or other learning-relevant constructs; the information contained in these models is usually used to drive the adaptivity in intelligent teaching systems [17]. In most adaptive systems, the learner model is hidden from the learner. However, open learner models (OLM) allow the learner to view the information that is contained within the system's model of the learner [8]. Making the learner model open to or allowing it to be viewed by the learner may increase learners' metacognitive skills, e.g. promote learner reflection, and help them to plan and monitor their learning [7].

OLMs allow learners to access learner model information through one or more visualisations, such as the very common skill meters [4,5,6,11,15,26], concept maps [15,22,30], hierarchal tree structures [15,19,21,22], networks [4,6], tree maps [3,4,6,21] word clouds [4,6], and radar plots [4,6,21]. Figure 1 shows some of the visualisations of the learner model from the Next-TELL [4] and LEA's Box OLMs [6]. Skill meters, at the top of Figure 1, indicate the level of knowledge by filling in the bar, and can be useful with a low number of topics. If the learner model has a larger number of topics, the user would need to scroll down to view all the topics. In contrast, the network visualisation that is below the skill meters in Figure 1, shows a larger number of topics in the same screen space, but it can be difficult to read if many nodes become very close together. The network uses different variables, such as size and colour, to indicate the level of understanding of the learner. (The larger and brighter the colour, the higher the level of understanding.) The radar plot that is below the network on the right of Figure 1 can also show the relative weaknesses or strengths of learner knowledge for different topics in a smaller space than the skill meters, but does not allow the domain structure to be shown. The tree map that is to the left of the radar plot can be useful when a large number of topics need to be shown because learner understanding of sub-topics can be explored by clicking on the parent topic. However, this means that users cannot compare topics from different parts of the tree. The level of understanding for each topic is indicated by the size of the corresponding rectangle. The topics in the word cloud at the bottom of Figure 1 are separated into two boxes: strong and weak topics. In the weak box, larger words indicate weaker skills, whereas those that are larger in the strong box are stronger skills.

The data in the learner model can come from the same system, as has traditionally been the case (e.g., [3,5,9,10,14,19,20,22,26]) or a variety of external sources (e.g., [4,6,25,28]). For example, in Next-TELL's

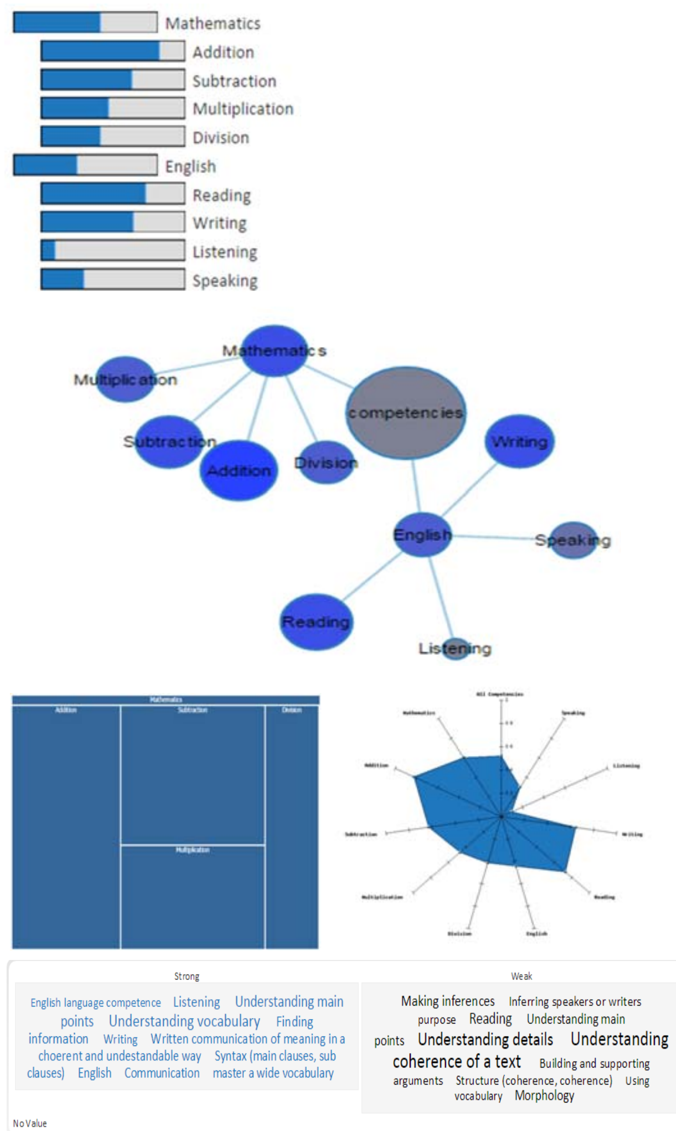


Figure 1: Examples of open learner model visualisations from Next-TELL OLM [4] and LEA's Box OLM [6]

OLM, data can come from different automated sources (e.g., quizzes, problems, virtual world activities) or manually entered sources (e.g., self-assessments, peer-assessments, or teacher assessments of the learner's skills) [4]. Using different data sources can allow different activities to be taken into account during the inference process that creates the learner model, much like portfolio assessment and e-portfolios use a variety of evidence when assessing learners [31,33]. However, using varied data sources may increase the likelihood of model uncertainty because of the variability in the data that is included.

Researchers whose focus is on managing uncertainty have recognised the problem of uncertainty within the learner modelling process [18]. Model uncertainty in general is based on the quality of the data that is influenced by different key components such as error, accuracy, consistency, completeness and precision [12]. To address these and other types of uncertainty, numerical techniques that account for uncertainty within the learner model have occasionally been used. These methods include Bayesian networks and fuzzy logic [18].

In this paper, we focus on uncertainty visualisation in open learner models in terms of inconsistency in the data over which the model reasons. For instance, a student may receive a low score on one quiz and score highly on all of the other quizzes; or there may be inconsistencies between automatically inferred data and self-assessments. Visualising uncertainty in the learner model can reveal these inconsistencies. Uncertainty in the learner model data can be indicated using different aspects of the visualisation (i.e., visual variables [13]). The use of well-selected visual variables can permit users to automatically identify the pattern depicted by those visual variables without having to focus their attention on this task [23]. Different methods of visually representing uncertainty within OLMs, using visual variables such as blur, opacity and arrangement have been proposed [13]. We here extend that work to measure the uncertainty, which is a precursor to visualising uncertainty in the learner model. Measuring uncertainty in our current research is based on identifying inconsistencies within the data set.

This paper is organised as follows. Section 2 presents work related to uncertainty in OLMs and Section 3 discusses uncertainty visualisation. Section 4 proposes how a teaching system can identify the uncertainty that is caused by inconsistency within a data set. Following this, Section 5 suggests the further benefit of uncertainty visualisation in OLMs that are jointly maintained by student and system.

2. UNCERTAINTY IN OLMs

The visualisation community has recently become increasingly aware of the importance of visualising the uncertainty that is present in data [2]. Understanding uncertainty in the data is important to allow users to make better decisions based on the information given in the OLM ([13]).

Unreliable data evidence can be obtained by students correctly guessing or accidentally making a mistake, both of which affect the state of the learner model [32]. Student modelling has sometimes bypassed the issue of model uncertainty, or handled it by using simple techniques, such as fuzzy logic; or complex techniques, such as Bayesian reasoning [18]. Fuzzy logic uses simple variables to represent the level of understanding for a learner, with imprecise values from within a range assigned to a variable to make it easier to understand and modify; Bayesian Networks, which are more complex, assign probability values to each node in the learner model representing the possibilities of different paths in a cause and effect relationship [18].

Different methods can be used to visualise the level of knowledge and the beliefs represented in the learner model. For example, the VisMod Bayesian Belief Network [37] is a learner model visualised as a concept map with nodes that relate to the level of understanding and links that indicate the learning sequences. The level of understanding is constructed based on the probability value within a particular node including the previous knowledge and the current data evidence. VisMod [37] uses different data sources to construct the learner model (self-assessment, teacher-assessment and evidence provided by the system). VisMod then manipulates different visual elements that include colour, size, proximity, line thickness and animation to show changes in the node at different time intervals and from different data sources. The student's and the system/teachers' beliefs are taken into account in the visualisation of the student model to indicate uncertainty that results from having these two sets of beliefs. Each belief is represented with a separate node using colour and size to indicate the strength of the level of understanding. The overall level of understanding of both beliefs are visualised as another node using the average of the two beliefs (the student and the system's beliefs about the student's understanding). The colour of the combined node comes from the belief (from either the system or student) that most influences the modelled level of student understanding.

Fuzzy logic is another way of dealing with uncertainty in OLMs. For example, the LOZ open learner model [27] uses vague linguistic values (strong, medium and weak) to represent the learner's level of knowledge. The learner model is used to select multiple choice questions, which are classified into three levels of difficulty (high, moderate and low), to learners based on their level of knowledge. When a student with a weak level of knowledge correctly answers an assessment task from the difficult level, the system indicates that there is uncertainty due to inconsistency between the two sources (level of knowledge and level of difficulty), so the system provides another question to avoid having a lucky guess or a slip unduly influence the learner model.

While these approaches provide information about uncertainty in the underlying model, learners and teachers could still benefit from viewing additional information about model uncertainty. However, this information can only be displayed once it has been measured, and to date little effort has been expended on quantifying or representing student model uncertainty with a view to visualising this information to the user.

As indicated above, visual variables can be used to represent uncertainty and communicate it to the user [24]. Figure 2 shows some visual variables that can be used in the context of this paper, with three levels of uncertainty indicated from left to right (low, medium and high) [24]. The uncertainty levels represented in Figure 2 can be applied to different types of OLM visualisations, for example, those presented in Figure 1.

Arrangement is used to indicate uncertainty, where messier arrangements show higher uncertainty [29]. Opacity can be used to show uncertainty by increasing the transparency of uncertain data [23] and blur can be used to represent uncertainty by increasing the fuzziness of the visual element with the uncertainty that is present in that element's underlying data [23]. The size or thickness of a dashed outline can indicate uncertainty: the thicker the dashed line, the higher the uncertainty [1].

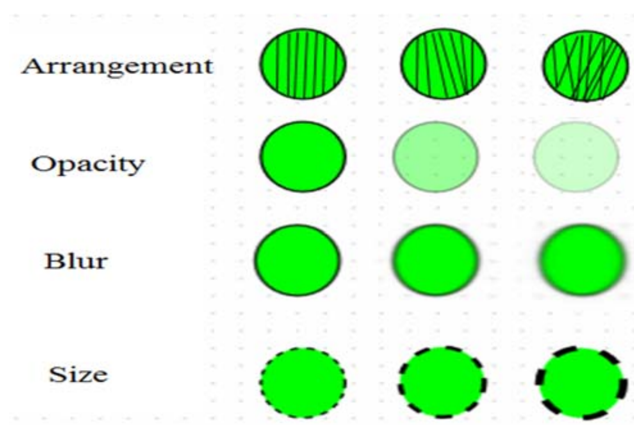


Figure 2: levels of uncertainty for visual variables.

3. EXAMPLES OF UNCERTAINTY VISUALISATION FOR OLMS

To take a step towards providing the learner with information about model uncertainty, we first demonstrate two uncertainty visualisations that have been integrated into the OLM of an existing teaching system: OLMlets [5]. OLMlets constructs a learner model using numerical weightings of student responses to multiple choice questions, with the learner model based on the last five questions that the learner has attempted in each topic. The age of the evidence affects its weight or influence on the model, with newer evidence being weighted more heavily. The learner model visualisation uses green to indicate correct knowledge and grey to indicate difficulty.

OLMlets has been extended to allow student self-assessments to be entered after their response to each multiple choice question (Figure 3). This additional source of learner model evidence complements the system's assessment of learner knowledge.

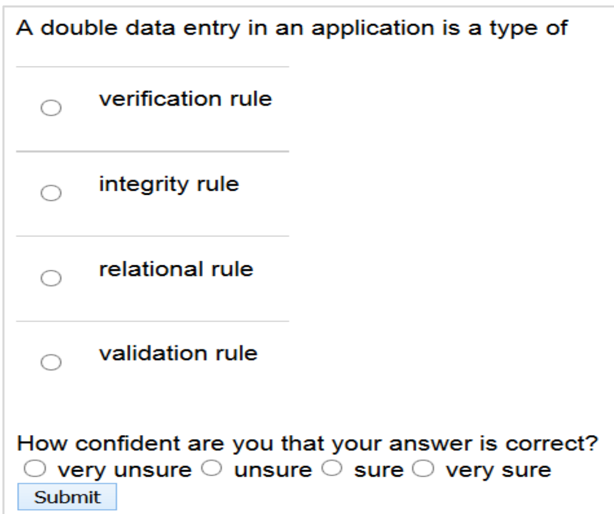


Figure 3: Question and self-assessment options.

OLMlets uses five visualisations to show the learner model [5]. In this paper, we focus on using the skill meters (similar to those in Figure 1) to visualise uncertainty because skill meters are commonly used in OLMs [8], and they are often popular when multiple visualisations are available (e.g. [4,5,15]). The first of our new OLM visualisations (Figure 4) shows skill meters placed side by side to represent the two models: the system's assessment of the learner and the student's self-assessment, with the skill meter fill (green) indicating level of understanding, and the remaining area of the skill meter (grey) showing the proportion of the topic in which the learner has difficulties. In this case, uncertainty (variability) can be seen in the discrepancy between the two models. The second approach (Figure 5) uses skill meters that combine the model that is based on the automatically inferred values (system model) with the model that is based on the student's self-assessments (student model) into a single set of skill meters. This version uses opacity (see Figure 2) to indicate where the two data sources conflict: the higher the transparency of a topic's green colour, the more inconsistent the data.

Like in other work that used confidence ratings [10, 20], the student model in the two visualisations is based in part on system inference, and in part on students selecting their level of confidence from a scale of 'very sure', 'sure', 'unsure', and 'very unsure' (see Figure 3). If the student selects 'very sure' or 'very unsure', this is interpreted to mean that the student is 100% confident about the correctness or incorrectness of their answer. If the student selects 'sure' as their confidence level, the system will weight the new evidence as 75% correct knowledge and 25% difficulty when visualising that information in the OLM. This is because the student believes more strongly that their answer is correct, but still acknowledges that they might be wrong. Selecting the 'unsure' option in the confidence level is represented as 75% difficulty and 25% correct knowledge.

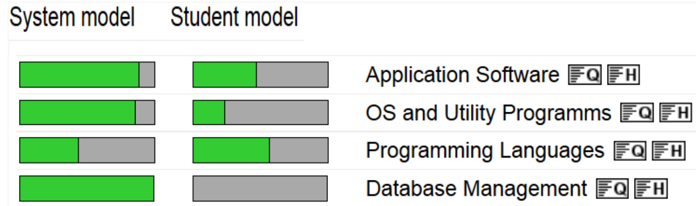


Figure 4: Uncertainty visualisation using two skill meters.

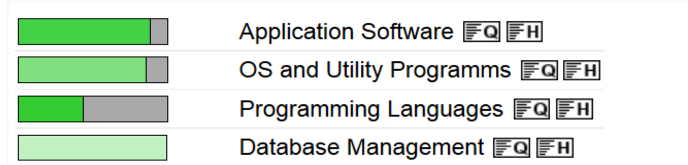


Figure 5: Uncertainty visualisation using opacity in skill meters.

Students viewing the two skill meters are given an indirect representation of model uncertainty that can be seen by comparing their beliefs to the system's beliefs about their level of knowledge (Figure. 4). Placing these models side by side should enable students to see the discrepancy between these two measures of their knowledge and enable them to recognise any inconsistency that is present. In the second approach to visualising uncertainty within an OLM (Figure 5), the opacity of the fill colour in the skill meter should similarly draw the learner's attention to topics where the data is inconsistent. This version indicates different uncertainty levels by increasing or decreasing skill meter opacity: database management is the least opaque topic and most uncertain, whereas the evidence used to infer the learner's knowledge of programming languages (system inference and student confidence ratings) is highly consistent which is why the skill meter is opaque.

To illustrate other visual variables (see Figure 2) for showing learner model uncertainty, we present OLM designs based on the Next-TELL [4] and LEA's Box [6] OLMs, shown in Figure 1.

Figure 6 shows how arrangement could be applied in skill meters, where an untidy arrangement in the skill meter fill (structure in Figure 6) indicates high uncertainty. Figure 6 also shows the hierarchy levels of topics and sub-topics within the underlying learner model (not present in the previous example of skill meters from OLMlets).

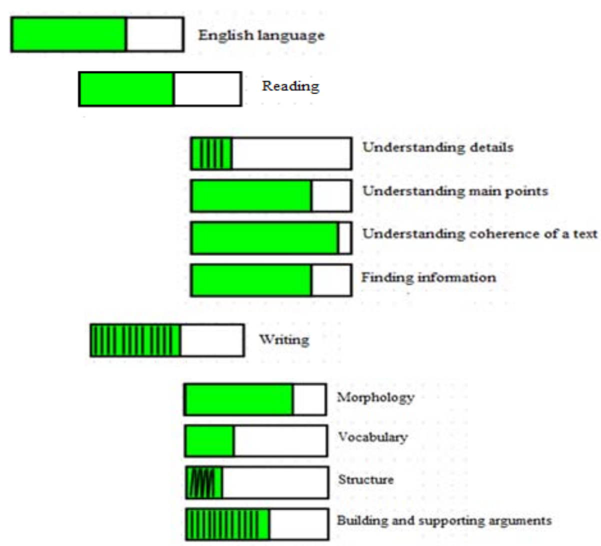


Figure 6: Uncertainty visualisation using arrangement in skill meters.

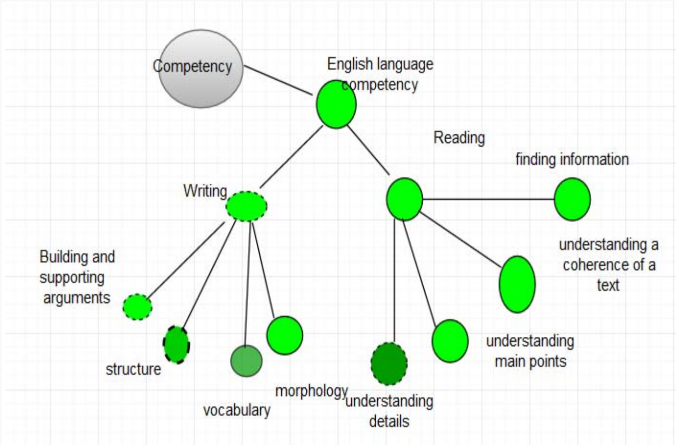


Figure 7: Uncertainty visualisation using a dashed line around nodes for uncertain topics in a network.

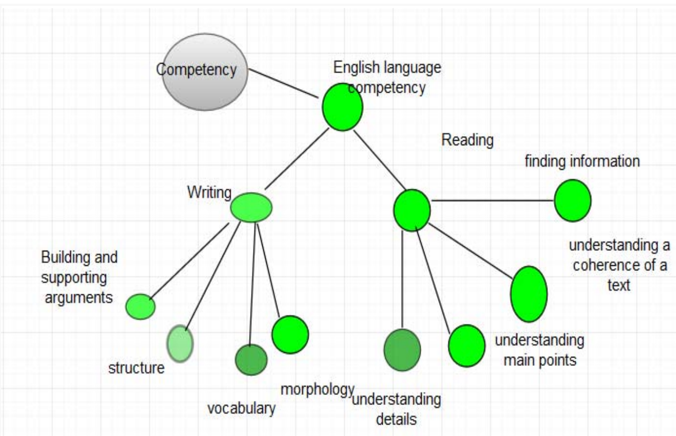


Figure 8: Uncertainty visualisation using opacity on nodes for uncertain topics in a network.

In addition to skill meters, Figure 1 showed network, radar plot, word cloud and tree map based visualisations in the Next-TELL [4] and LEA's Box [6] OLMs. The network visualisation uses size and colour to indicate the knowledge level of the topic. Larger and brighter nodes indicate that the learner has achieved a higher knowledge or competency level for that topic. Using a dashed line around the edge of the node could indicate whether there is uncertainty associated with that topic's assessment, and using different levels, indicated by the size (thickness) of dashed lines, can illustrate the uncertainty level (Figure 7). Furthermore, uncertainty in the sub-topics can be inherited by the parent topic (as can also occur with skill meters).

Figure 7 includes two of the main topics (reading and writing) that are subtopics of English language. These two sub-topics also have several sub-topics of their own. The writing topic has one sub-topic (building and supporting arguments) that shows a low level of uncertainty by a thin dashed outline, and one sub-topic (structure) that has medium level of uncertainty, shown by a thicker dashed outline. The other two sub-topics do not contain uncertainty or conflicting data. The parent topic (writing) takes the average of the uncertainty levels that are associated with each of its sub-topics. The parent topic is visualised with a low level of uncertainty that is the result of the uncertainty that it has inherited from its children by calculating the average of the uncertainty weight for all sub-topics (1 had low uncertainty, 1 had medium uncertainty, and 2 had no uncertainty). Instead of using dashed lines, the same information could be represented using opacity (see Figure 8).

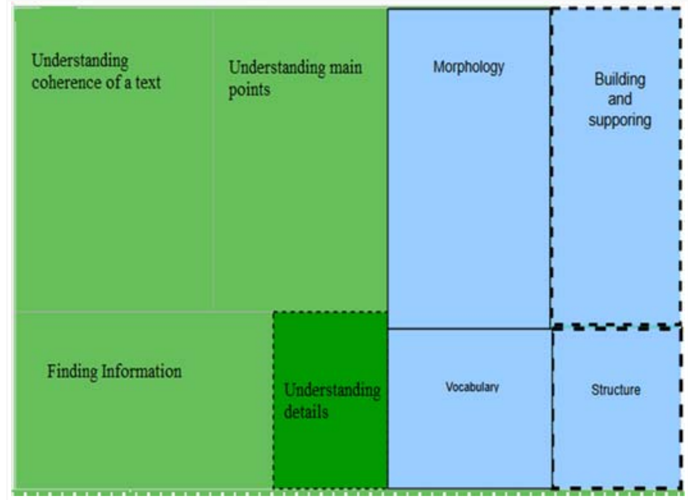


Figure 9: Uncertainty visualisation using the size of dashed line in tree map.

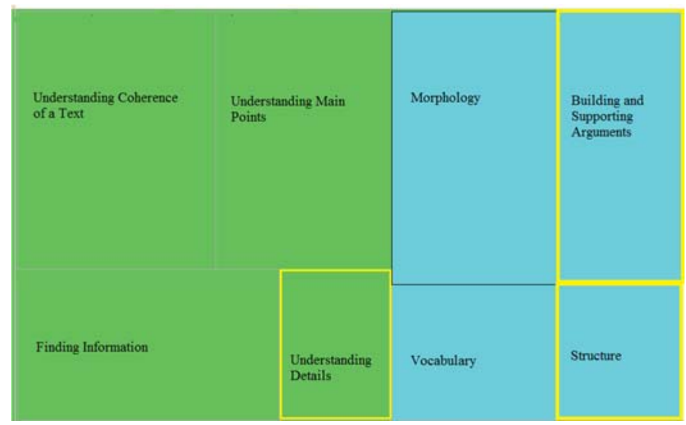


Figure 10: Uncertainty visualisation using line colour in tree map.

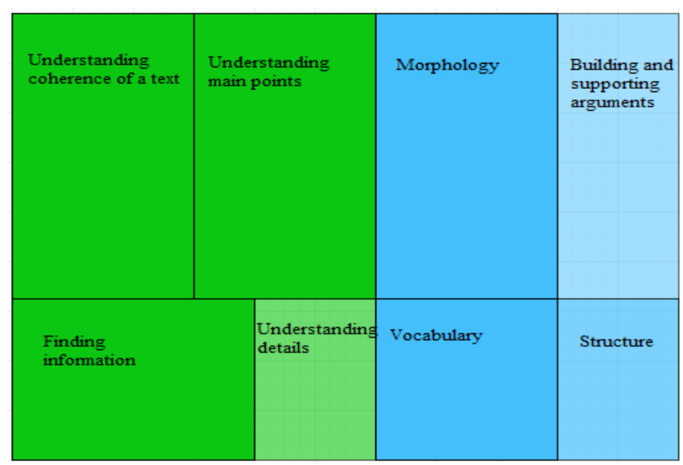


Figure 11: Uncertainty visualisation using opacity of the colour in tree map.

When a large number of topics or competencies are contained in the learner model, tree maps may be useful to allow learners to explore different levels of a hierarchically structured learner model [3,4]. Both brightness and line colour have been used as an indicator of uncertainty in tree maps in the field of simulation and visualisation [16]. Following from these efforts, we propose using a dashed line around the topic border to represent uncertainty within a tree map (Figure 9), where different levels of size (thickness) of the dashed line indicate the uncertainty level. This can also be done by varying the brightness and colour of the line around the edge of a model topic (Figure 10) or through the use of opacity (Figure 11).

Figure 12: Uncertainty visualisation using the blur in word cloud.

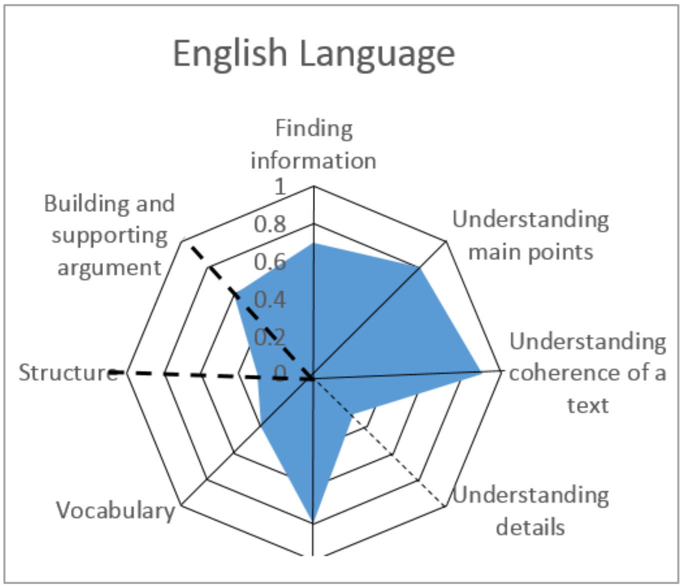


Figure 13. Uncertainty visualisation using the size of dashed line in radar plot.

Word clouds allow people to quickly identify stronger topics because the text is larger (and, in the case of the Next-TELL [4] and LEA's Box [6] OLMs, also the weaker competencies in the second word cloud (see Figure 1). To show uncertainty, blur could be applied to the text: the fuzzier the text, the higher the uncertainty (structure in Figure 12). Colour could also be used to help indicate the grouping of sub-topics. Figure 12 shows two groups of sub-topics where each group has its own colour (orange or blue), to allow some structuring of the domain, otherwise difficult to achieve with word clouds.

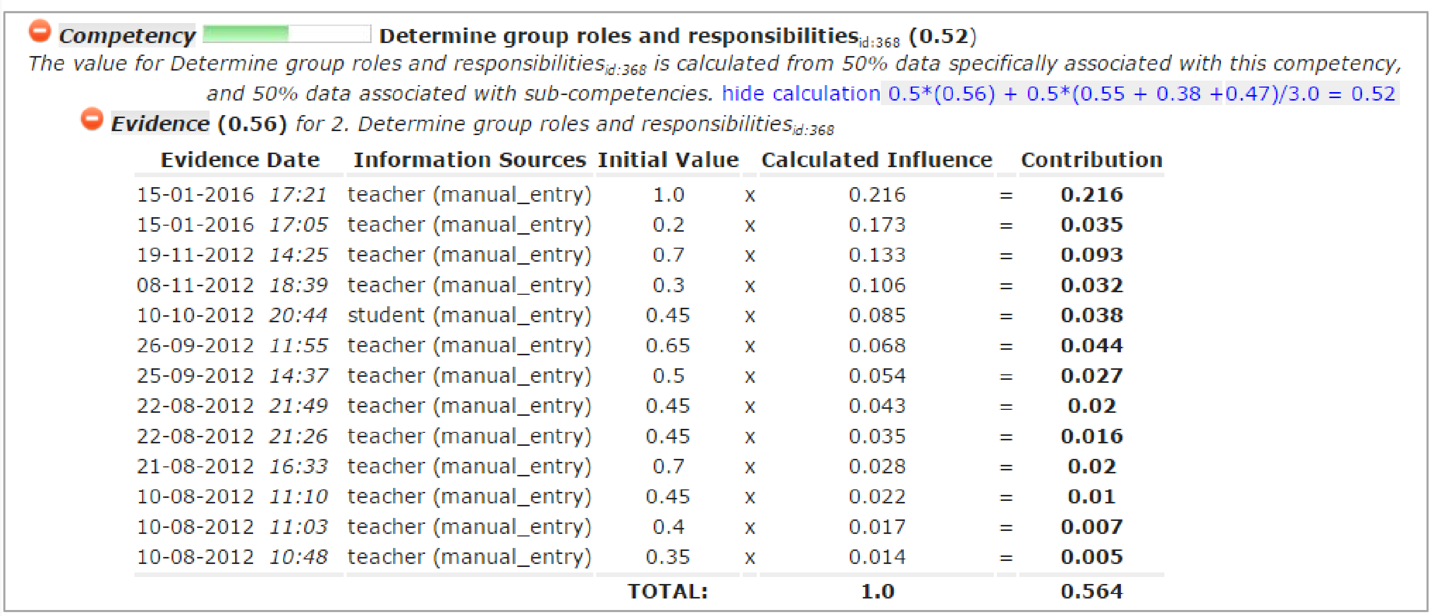
As illustrated in Figure 13, radar plots can show uncertainty by using, for example, a dashed line assigned to a topic with uncertainty in the data associated with it (as previously illustrated for the network and tree map visualisations).

This section has presented several visualisation techniques that could be used to display uncertainty within open learner models, showing different levels of uncertainty using the visual variables of arrangement, opacity, blur and size (line thickness); and two separate versions of the learner model placed side-by-side in the simpler skill meter visualisation. As indicated in the introduction, OLMs may facilitate learner reflection, planning and self-monitoring, which can be a powerful way to help promote effective independent learning [7]. However, for this to be effective, some understanding of the level of uncertainty in the underlying model is needed to enable learners to better understand the accuracy of that data, and so better use the learner model information when making decisions about their learning. The next section proposes a method to measure uncertainty.

4. MEASURING UNCERTAINTY USING MULTIPLE DATA SOURCES

Rather than managing and designing around uncertainty, we want to measure uncertainty in the learner model within a data set and communicate that uncertainty. To measure uncertainty, we should first understand how the data are used within the learner model based on the modelling process that is used within a particular system. While there are many learner modelling techniques (see e.g. [17,18]) for the example in this paper we focus on measuring uncertainty in models that use a numerical weighting method. In the Next-TELL [4] and LEA's Box [6] OLMs, the data can come from several (or many) different data sources, and all evidence is used when calculating learner model values. However, each piece of evidence may influence the learner model differently, with all of the corresponding weights for each topic in the learner model summing to 1 [4].

In the Next-TELL and LEA's Box OLMs, teachers can configure the weight of different types of evidence. For example, the teacher may assign a higher weight to automated assessment sources than the manually entered data that is collected through self or peer assessments. The level of influence for each data set is normalised so that they sum to 1.0. The value of the data (v), where v is greater than or equal to 0.0 and v is less



Competency **Determine group roles and responsibilities$_{id:368}$ (0.52)**

The value for Determine group roles and responsibilities$_{id:368}$ is calculated from 50% data specifically associated with this competency, and 50% data associated with sub-competencies. hide calculation 0.5*(0.56) + 0.5*(0.55 + 0.38 +0.47)/3.0 = 0.52

***Evidence* (0.56)** *for 2. Determine group roles and responsibilities$_{id:368}$*

Evidence Date	Information Sources	Initial Value		Calculated Influence		Contribution
15-01-2016 *17:21*	teacher (manual_entry)	1.0	x	0.216	=	**0.216**
15-01-2016 *17:05*	teacher (manual_entry)	0.2	x	0.173	=	**0.035**
19-11-2012 *14:25*	teacher (manual_entry)	0.7	x	0.133	=	**0.093**
08-11-2012 *18:39*	teacher (manual_entry)	0.3	x	0.106	=	**0.032**
10-10-2012 *20:44*	student (manual_entry)	0.45	x	0.085	=	**0.038**
26-09-2012 *11:55*	teacher (manual_entry)	0.65	x	0.068	=	**0.044**
25-09-2012 *14:37*	teacher (manual_entry)	0.5	x	0.054	=	**0.027**
22-08-2012 *21:49*	teacher (manual_entry)	0.45	x	0.043	=	**0.02**
22-08-2012 *21:26*	teacher (manual_entry)	0.45	x	0.035	=	**0.016**
21-08-2012 *16:33*	teacher (manual_entry)	0.7	x	0.028	=	**0.02**
10-08-2012 *11:10*	teacher (manual_entry)	0.45	x	0.022	=	**0.01**
10-08-2012 *11:03*	teacher (manual_entry)	0.4	x	0.017	=	**0.007**
10-08-2012 *10:48*	teacher (manual_entry)	0.35	x	0.014	=	**0.005**
		TOTAL:		**1.0**		**0.564**

Figure 14: Next-TELL learner model calculation evidence screen showing the calculation of a student's competency level for group roles and responsibilities [4].

Table 1: Example of uncertainty calculations and weighting when an outlier (shown in *italics*) is present in the data evidence.

Information Source	Initial Value	Calculated Influence (on knowledge)	Uncertainty Weight
Self-Assessment	*.2*	*.345*	*.345*
Peer-Assessment	.9	.243	0
Teacher-Assessment	.8	.174	0
Quiz1	.7	.152	0
Quiz2	.9	.086	0
	Total:	1.00	.345

than or equal to 1.0, is then multiplied by the level of influence to show how much that piece of evidence contributes to the learner model.

In the Next-TELL OLM, the learner model calculation can be viewed by the teacher and the student (Figure 14) [4]. Like with the OLMlets example (Figure 4), described in Section 3, users can see the inconsistent data when viewing the screen that shows the model calculation (Figure 14), but they only see this inconsistency if they invest additional effort to search through the data evidence. This effort requires them to look at each line in the whole calculation and compare those lines to one another. Taking advantage of visual communication channels to show the uncertainty in the data upon which the learner model is based could help learners to identify inconsistencies without going through all of these calculations, which holds the potential to better support their self-regulation and planning activities. As indicated above, this was achieved in a quite simple way when extending the OLMlets skill meters to take account of two sources of data (system and student assessments of the student's knowledge). We propose the following approach where there may be more complex relationships between data from different activities or different parts of activities or, indeed, from different data sources as in the Next-TELL [4] and LEA's Box [6] OLMs.

In order to visualise uncertainty based on inconsistency in the underlying data, the source and the level of the uncertainty must be measured. Knowing the source of the uncertain data helps us to indicate the level of uncertainty in the learner model by summing the influence weight for all the sources that contribute to model uncertainty. To identify the source of the inconsistent data, we apply outlier analysis to detect inconsistencies in the data. Outliers are based on the concept of boxplots. To detect outliers, formula (1) and (2) are used to calculate the upper fence and lower fence. These fences are based on the data's inter-quartile range (IQR), which is the difference between the first (q1) and third quartiles (q3), with the data that are outside these fences classified as outliers [36].

$$Upper\ Fence = q3 + 1.5(IQR) \qquad (1)$$

$$Lower\ Fence = q1 - 1.5(IQR) \qquad (2)$$

Considering the example shown in Table 1, the learner model has five data sources contributing to the calculation of learner knowledge or competency, and each source has its initial score value and an associated weighting. Applying formula (1) and (2) to the data in Table 1 results in an upper fence of 1.2 and a lower fence of 0.4. In Table 1, the self-assessment scores are outside this range (i.e., they are outliers). The uncertainty level can now be measured by detecting how much weight is assigned to each outlier. Summing all the weights from all of the outliers provides the value for the model's uncertainty weight, which indicates how much these uncertain pieces of data influence the model. From Table 1, this is .345.

Similar to the learner modelling process, new pieces of evidence that are classified as outliers influence the level of uncertainty associated with that model attribute more than an old piece of evidence would. Like the weights that are associated with topics, uncertainty values range from 0 (no uncertainty) to 1 (high uncertainty). However, using real numbers communicates a level of precision that is not present within the system. As a result, this range is subdivided into three levels of uncertainty: namely, low (0-0.3), medium (0.3-0.7) and high (0.7-1.0). These three levels can be visualised using the variables shown in Figure 2, and illustrated in Section 3. Since there is only one outlier detected from the information given in Table 1 and it has a weight of 0.345, a medium level of uncertainty is associated with that competency. The ability to determine the amount of uncertainty that is associated with a specific competency allows us to show that uncertainty to users so that learners or teachers can use this information to support their planning and decision-making tasks, facilitating some of the metacognitive benefits argued for OLMs [7].

5. UNCERTAINTY VISUALISATION FOR LEARNER MODELS JOINTLY MAINTAINED BY STUDENT AND SYSTEM

Beyond supporting learner planning and decision making as argued previously, visualising learner model uncertainty may be useful to learners when they are using interactively maintained learner models. These types of OLMs include those that allow the learner to try to persuade the teaching system to change learner model values because the learner disagrees with some aspect of the system's model. This can be valid, for example, if a student has done some reading, exercises, etc., away from the teaching system; or if they had achieved correct answers through (partial) guessing. This challenge to the system's model can succeed by having learners verify their proposed change through responses to additional questions or assessment items that are administered by the system (e.g. [9, 35]); or by having learners negotiate a change to the learner model through a two-way discussion of the learner model content. This discussion takes place between the learner and the system with the goal of having both parties agree on the model (e.g., [10,14,20]), but keeping separate representations if agreement is not achieved. Both these approaches to interactively maintained learner models (persuadable and negotiated), as well as aiming for a more accurate learner model, also aim to prompt reflection (as described above), through the process of challenging and discussing the model. In cases where students can challenge the system's model, as described above, an indication of the certainty of data could be highly beneficial, to focus updates onto topics with the most uncertain or inconsistent data, therby making the learner model more accurate and improving subsequent adaptation. This is a timely topic as current projects (in the areas of persuadable [6] and negotiated [34] learner models) strive to involve the learner more in the modelling process.

6. SUMMARY

Building on the work of [13], this paper proposes several approaches to uncertainty visualisation using different methods such as the width of a dashed line, opacity of OLM elements, the application of blur and arrangement of visual elements within a learner model component. The visual presentation of model uncertainty is based on inconsistency in the underlying model's data. The ability to see model uncertainty was integrated into the OLMlets system through two visualisations that are based on the commonly used skill meter representation of learner knowledge. These visualisations are being used in an ongoing study that investigates the effect of uncertainty visualisation on students' self-assessments and learning outcomes.

In addition to this work, a method for identifying inconsistencies in the underlying learner model was developed, and was described with reference to the Next-TELL [4] and LEA's Box [6] OLMs, which have potentially many data sources. This method uses outlier analysis to identify data that contribute to model uncertainty. The identified data is then assigned a weight based on the underlying learner modelling formula. This information is used to determine the level of uncertainty that is present in different model attributes so that the uncertainty can be visualised as proposed; the proposed OLM visualisations used the network, tree map, word cloud and radar plot versions of the OLM from the Next-TELL and LEA's Box OLMs. These visualisations, which rely on outlier analysis for identifying uncertainty, will be integrated into an OLM as the next step towards supporting metacognitive activities and learner model negotiation or persuasion in interactively maintained learner models.

7. ACKNOWLEDGEMENT

The first author is supported by a PhD Scholarship from the Ministry of Higher Education in Oman. The LEA's Box project is supported by the European Commission (EC) under the Information Society Technology priority FP7 for R&D, contract 619762 LEA's Box, building on contract 258114 Next-TELL. This document does not represent the opinion of the EC and the EC is not responsible for any use that might be made of its contents.

8. REFERENCES

1. Bertin, J. *Semiology of Graphics: Diagrams, Networks, Maps*. University of Wisconsin Press, Madison, WI, 1983.
2. Bonneau, G.P., Hege, H.C., Johnson, C.R., Oliveira, M.M., Potter, K., Rheingans, P. and Schultz, T. Overview and state-of-the-art of uncertainty visualization. In Hansen, C.D., Chen, M., Johnson, C.R., Kaufman, A.E. and Hagen, H. (eds.), *Scientific Visualization, Uncertainty, Multifield, Biomedical, and Scalable Visualization* Springer-Verlag, London, 3-27, 2014.
3. Brusilovsky, P., Hsaio, I.-H., and Folajimi, Y. QuizMap: open social student modeling and adaptive navigation support with treemaps. In Kloos, C.D. Gillet, D., Crespo Garcia, R.M., Wild, F., and Wolpers, M. (eds.), *Towards Ubiquitous Learning*, EC-TEL, Springer-Verlag, Berlin, 71-82, 2011.
4. Bull, S., Johnson, M.D., Masci, D. and Biel, C. Integrating and visualising diagnostic information for the benefit of learning. In Reimann, P., Bull, S., Kickmeier-Rust, M., Vatrapu, R.K. and Wasson, B. (eds.), *Measuring and Visualizing Learning in the Information-Rich Classroom*, Routledge/Taylor and Francis, 167-180, 2016.
5. Bull, S. and Mabbott, A. 20,000 inspections of a domain-independent open learner model with individual and comparison views. In Ikeda, M., Ashley, K. and Chan, T-W. (eds.) *Intelligent Tutoring Systems*, Springer-Verlag, Berlin Heidelberg, 422-432, 2006.
6. Bull, S., Ginon, B., Boscolo, C. and Johnson, M.D. Introduction of learning visualisations and metacognitive support in a persuadable open learner model, In Gasevic, D. and Lynch, G. (eds.) *Learning Analytics and Knowledge 2016*, ACM, in press.
7. Bull S., and Kay, J. Open learner models as drivers for metacognitive processes. In Azevedo, R., Aleven, V. (eds.), *International Handbook on Metacognition and Learning Technologies*, Springer, New York, 349-366, 2013.
8. Bull, S. and Kay, J. 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), 293–331, 2016.
9. Bull, S., Mabbott, A. and Abu-Issa, A. UMPTEEN: Named and anonymous learner model access for instructors and peers. *International Journal of Artificial Intelligence in Education, 17* (3), 227-253, 2007.
10. Bull, S., Pain, H. "Did I say what I think I said, and do you agree with me?": Inspecting and questioning the student model. in Greer, J. (ed.), *Proceedings of World Conference on Artificial Intelligence and Education,* Association for the Advancement of Computing in Education, Charlottesville VA, USA, 501-508, 1995.
11. Corbett, A.T., and Bhatnagar, A. Student modeling in the ACT programming tutor: adjusting a procedural learning model with declarative knowledge. In Jameson, A., Paris, C., and Tasso, C. (eds.), *User modeling,* Springer, New York, 243-254, 1997.
12. Correa, C.D., Chan, Y.H., and Ma, K.L. A framework for uncertainty-aware visual analytics. *VAST 09 - IEEE Symposium on Visual Analytics Science and Technology, Proceedings,* 51-58, 2009.
13. Demmans Epp, C., and Bull, S. Uncertainty representation in visualizations of learning analytics for learners: Current approaches and opportunities. *IEEE Transactions on Learning Technologies*, 8 (3), 242-260, 2015.
14. Dimitrova, V. StyLE-OLM: Interactive open learner modelling. *International Journal of Artificial Intelligence in Education, 13* (1), 35-78, 2003.
15. Duan, D., Mitrovic, A. and Churcher, N. Evaluating the effectiveness of multiple open student models in EER-Tutor. In Wong, S.L., Kong, S.C. and Yu, F-Y. (eds). *International Conference on Computers in Education,* Asia-Pacafic Society for Computers in Education, 86–88, 2010.
16. Griethe H. and Schumann, H. The visualization of uncertain data: methods and problems. In *Proceedings of simulation and visualization*, SCS Publishing House, Magdeburg, Germany, 143-156, 2006.
17. Holt, P., Dubs, S., Jones, M. and Greer J. The state of student modeling. In Greer, J. and McCalla, G. (eds.), *Student Modeling: The Key to Individualized Knowledge-Based Instruction*, Springer, New York, 3-39, 1994.
18. Jameson, A. Numerical uncertainty management in user and student modeling: An overview of systems and issues. *User Modeling and User-Adapted Interaction*, *5(3),* 193-251, 1995.
19. Kay, J. Learner know thyself: Student models to give learner control and responsibility. In Halim,Z., Ottomann,T., Razak,Z. (eds). *International Conference on Computers in Education*, 17-24, AACE, 1997.
20. Kerly, A., Ellis, R. and Bull, S. CALM system: A conversational agent for learner modelling. In Ellis, R., Allen, T. and Petridis, M. (eds.), *Applications and Innovations in Intelligent Systems,* XV, *Proceedings of AI-2007, 27th SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence,* Springer Verlag, 89-102, 2007.
21. Mathews, M., Mitrovic, A., Lin, B. Holland, J. and Churcher, N. Do your eyes give it away? Using eye tracking data to understand students' attitudes towards open student model representations. In Cerri, S.A., Clancey,W.J., Papadourakis,G. and Panourgia,K. (eds). *Intelligent Tutoring Systems*, Springer, Heidelberg, 422-427, 2012.
22. Mabbott, A. and Bull, S. Alternative views on knowledge: Presentation of open learner models. In Lester, J.C., Vicari, R.M. and Paraguacu, F. (eds.), *Intelligent Tutoring Systems,* Springer-Verlag, Berlin, Heidelberg, 689-698, 2004.
23. MacEachren, A.M. Visualizing uncertain information. *Cartographic Perspective, 13,* 10-19, 1992.
24. MacEachren, A.M., Roth, R.E., O'Brien, J., Swingley, D. and Gahegan, M. Visual semiotics and uncertainty visualization: an empirical study. *IEEE Trans., Visualization and Computer Graphics*, 18 (12), 2496–2505, 2012.
25. Mazzola, L. and Mazza, R.: GVIS: A Facility for adaptively mashing up and presenting open learner models. In Wolpers, M., Kirschner, P.A., Scheffel, M., Lindstaedt, S. and Dimitrova, V. (eds.), *EC-TEL 2010,* Springer-Verlag, Berlin Heidelberg, 554-559, 2010.
26. Mitrovic, A. and Martin, B. Evaluating the effect of open student models on self-assessment. *International Journal of Artificial*

Intelligence in Education, 17 (2):121–144, 2007.

27. Mohanarajah, S., Kemp, R. and Kemp, E. Opening a fuzzy learner model. In Proceedings of the Workshop on Learner Modelling for Reflection, *International Conference on Artificial Intelligence in Education*, Amsterdam, Netherlands, 62-71, 2005.

28. Morales, R., Van Labeke, N., Brna, P. and Chan, M. Open learner modelling as the keystone of the next generation of adaptive learning environments. In Mourlas, C. and Germanakos, P. (eds.), *Intelligent User Interfaces: Adaptation and Personalization Systems and Technologies*. Information Science Reference, ICT Global, London, 288-312, 2009.

29. Morrison, J.L. A theoretical framework for cartographic generalization with the emphasis on the process of symbolization. *International Yearbook of Cartography*, 14, 115-127, 1974. Cited in [24].

30. Perez-Marin, D., Alfonseca, E., Rodriguez, P. and Pascual-Neito, I. A study on the possibility of automatically estimating the confidence value of students' knowledge in generated conceptual models. *Journal of Computers, 2* (5), 17-26, 2007.

31. Raybourn, E.M. and Regan, D. Exploring e-portfolios and independent open learner models: Toward army learning concepts 2015. Interservice / Industry Training, Simulation and Education Conference Proccedings, Florida, USA, 2011.

32. Reye J. Student modelling based on belief networks. *International Journal of Artificial. Intelligent in Education,* 14(1), 63-96, 2004.

33. Seers, M., Gijbels, D. and Thurlings, M. The relationship between students' perceptions of portfolio assessment practice and their approaches to learning. *Educational Studies, 34* (1), 35-44, 2008.

34. Suleman, R.M., Mizoguchi, R. and Ikeda, M. Negotiation-driven learning. In Conati, C., Heffernan, N., Mitrovic, A. and Verdejo, M.F. (eds.), Artificial Intelligence in Education, Springer International Publishing, 470-479, 2015.

35. Thomson, D. and Mitrovic, A. Preliminary evaluation of a negotiable student model in a constraint-based ITS. *Research and Practice in Technology Enhanced Learning, 5* (1), 19-33, 2010.

36. Tukey, J.W. *Exploratory Data Analysis.* Addison-Wesley, Reading, MA, USA, 1977.

37. Zapata-Rivera, J.D., and Greer, J.E. Interacting with inspectable Bayesian student models. *International Journal of Artificial Intelligence in Education, IOS Press, 14(2),* 127–163, 2004.