2018 Australasian Software Engineering Conference, 26-30 Nov, Adelaide, Australia (c) 2018 IEEE

Evaluating an open learner model visualisation prototype tool with User eXperience metrics

Check-Yee Law
Faculty of Information Science and Technology
Multimedia University
Melaka, Malaysia
cylaw@mmu.edu.my

John Grundy
School of Information Technology
Monash University
Melbourne, Australia
John.Grundy@monash.edu

Andrew Cain
School of Information Technology
Deakin University
Melbourne, Australia
andrew.cain@deakin.edu.au

Karola von Baggo
Faculty of Science, Engineering and Technology
Swinburne University of Technology
Melbourne, Australia
kvonbaggo@swin.edu.au

Rajesh Vasa
Applied Artificial Intelligence Institute
Deakin University
Geelong, Australia
rajesh.vasa@deakin.edu.au

***Abstract*—Learning management systems (LMS) are widely used in many educational environments. Although usability is widely acknowledged as a desired quality attribute, it has not been deeply investigated in the context of LMS systems where it has to align with pedagogy. In this research we evaluate extensions to an LMS designed to support the implementation of constructive alignment for technical units using a Task-Oriented Portfolio approach. The LMS application was extended with a range of open learner model-based visualisations to help staff and students monitor progress towards achieving unit intended learning outcomes throughout unit delivery. This paper reports on a case study of a formative usability testing of the prototype tool by using User eXperience (UX) metrics. The results unveil the potential usability issues and ways to improve the design of the prototype tool to better cater the actual target user needs.**

Keywords—software development; educational software; prototype tool; outcome-based learning; constructive alignment; open learner model; visualisation; formative usability testing; user experience metrics; Pearson correlation

I. INTRODUCTION

Software and product development activities are heavily influenced by the needs and expectations of end-users. This means that the user is a critical factor and user-centred development processes have brought tremendous benefits to software and product development processes [1]. Understanding potential user needs, requirements and experience is core to the user-centred process of developing any new technological tools. Additionally, understanding users is a learning process and acknowledging this is critical for better development, implementation, and testing of a software product.

Performing a formative usability testing is a well-accepted method for evaluating user experience and multiple tests are often needed to ensure an effective outcome. User experience is defined in ISO 9241-210 as "*a person's perceptions and responses that result from the use or anticipated use of a product, system or service*" [2]. This points out how important it is to adopt user perspective into any software development process. To do this, User eXperience (UX) metrics are used to convey information about the effectiveness and efficiency of a product as well as user satisfaction. UX metrics provide software development teams with in-depth insights on the drawbacks and problems, and the improvement potentials to better focus the development of the product to the needs and goals of users.

Various software systems and applications have been widely used to support teaching and learning. For example, the use of learning management systems (LMS) for delivering, documenting, tracking and managing training programmes or education courses [3] and e-portfolios for students to document and showcase their learning artefacts produced over their academic life [4]. Although some existing learning management systems allow teachers to set up a link between the learning activities and institutional missions and/or program goals, the linking is restricted to fulfilling the institutional reporting purposes [5] and has little support to facilitate student self-regulated learning. Extending LMSs with open learner visualisations that show progress toward achieving the intended learning outcomes (LOs) is posited to help motivate students in taking a greater ownership of their learning in achieving the intended LOs.

Doubtfire is an LMS developed to support frequent formative feedback in Task-Oriented Portfolio assessment [6]. In Doubtfire, students are able to view unit tasks and submit their work for feedback. Staff are then able to provide formative feedback and sign tasks off by indicating the task status. In supporting an outcome-based student-centred learning environment, Doubtfire system has been enhanced with novel open learner model visualisations to provide direct support for indicating the links between assessment tasks and the unit learning outcomes (LOs), and supporting student self-reflection on their learning to encourage student self-regulated learning – this enhancement is Doubtfire++.

This paper reports on a case study carried out to evaluate Doubtfire++ from a user experience perspective to investigate the effectiveness and efficiency of the interface that supports an outcome-based, student-centred teaching approach. We describe how we designed a set of usability tests in the context of teaching philosophy, in this case, with Doubtfire++ and key learnings from these.

The rest of this paper is structured as follows. Section II reviews literatures and research work related to the use of the teaching approach and tools in supporting student-centred learning environment that provide background and motivation to this study. Section III illustrates the interface design of Doubtfire++. This is followed by a case study of the usability testing that includes research methodology and data collection process in Section IV. We report the results obtained in Section V. A detailed discussion of the findings is presented in Section VI. We then conclude with a summary and highlight key future research directions in Section VII.

II. Background And Motivation

A. Application of the Constructive Alignment (CA) Model

To describe a student-centred learning environment, Biggs proposed the Constructive Alignment (CA) model [7]. Constructive refers to a situation where learners learn by doing or interacting with activities to construct their own knowledge whereas alignment means both teaching activities and assessment that are aligned to the intended learning outcomes (ILOs). Biggs articulated that the teaching and learning activities underpin active student engagement and strongly recommended portfolio assessment.

The constructive alignment approach has been adopted by Cain in Task-Oriented Portfolio teaching for programming related units, with an additional focus on frequent formative feedback [6]. This means assessment tasks are decomposed into small, frequent tasks that firmly link to the intended learning outcomes to guide student learning. In so doing, formative feedback becomes an essential part in this model for staff to provide iterative feedback to support student active construction and gain of knowledge. Doubtfire was designed and developed to support frequent formative feedback and serve as a communication platform for both staff and students.

B. Doubtfire

Doubtfire, as illustrated in Fig. 1, is a web application designed to support the frequent formative feedback cycles of the Task-Oriented Portfolio teaching approach [8]. While teaching staff use it to outline assessment tasks and provide feedback to students, it is used by students to keep track of their progress [8]. Within Doubtfire, students can view the unit assessment tasks and submit work for feedback. Staff are then able to provide formative feedback and sign tasks off as complete, or require students to resubmit work if required. Students can track their learning progress through *task list* and *burndown chart*.

Task list (Fig. 2) exploits different colours for students to track the status of each task. It provides an overview of the status of all tasks and support student engagement with frequent formative feedback on their tasks. The task status (i.e., Redo, Resubmit, Discuss, Complete) is encoded by a range of colour hues to show the status of a task such as brown for 'Redo', yellow for 'Resubmit', blue for 'Discuss' and green for 'Complete'. Students can easily identify the task that needs their immediate attention.

The task progress can be visualised by using a burndown chart (Fig. 3). A line connects the total number of tasks completed in a particular week with the total number of tasks completed in the following week throughout the semester. This line forms the slope and explicitly shows the relationship from one week to another throughout the whole semester. A flat slope indicates that a student does not make any progress for that week whereas a steeper slope denotes that more tasks have been completed, indicating good progress. Through the slope formed within a timeframe, it gives a stronger implication of trend relationships that show how a student manages and approaches the assessment tasks throughout the semester. This chart is augmented by colour to show students' actual progression, target progression and projected progression. It keeps students aware of the number of tasks remaining at a certain point of time. Students can also estimate effort required to complete their tasks to meet the minimum requirement of their desired grade. Doubtfire has been used in the teaching of programming units to address issues regarding time management [8]. To better support student self-reflection on their achievement in terms of learning outcomes, we have enhanced Doubtfire with open learner model visualisations.

C. Open Learner Models (OLM)

The open learner modeling concept, proposed by Self [9], has been extensively being applied in the development of OLM visualisation tools. Such tools have been used in various courses to present a learner's learner model that contains learning progress data in various computer-based representational formats. Exposing students to their own learner models is believed to bring immense educational benefits to them as well as the teaching staff. OLM tools have been shown to improve learner meta-cognitive activities, including self-assessment, self-regulation, self-reflection and help users to take greater control over the learning process [10]. It facilitates student self-reflection [11, 12], improves student engagement in the learning process [13-15] as well as encourage collaborative learning [16, 17] and hence helps students to acquire meta-cognitive skills [12, 18].

Developing an effective interface for presenting the learner model and for supporting user interaction is one of the key challenges in OLM research [16]. From our previous study, we proposed a range of open learner model representational formats to display student learning task status and achievement in terms of learning outcomes [19]. We have identified a set of OLM visualisations that are likely to be useful and accepted by our stakeholders through a detailed survey and interviews [20]. We see great educational benefits in enhancing Doubtfire with open learner model visualisations in supporting Task-Oriented Portfolio teaching in student-centred learning environment.

III. Doubtfire++

While Doubtfire was designed to support teaching of units using frequent formative feedback and constructive alignment, there is no direct support for indicating which tasks related to which learning outcomes, from either a staff or student perspective. To more effectively address potential misalignment issues and to better facilitate deep student learning through self-reflection to encourage student self-regulated learning, we have enhanced Doubtfire with visualisations adopted from open learner model research to allow staff and student to explore the links between tasks and learning outcomes, and to monitor student progress toward achievement of learning outcomes.

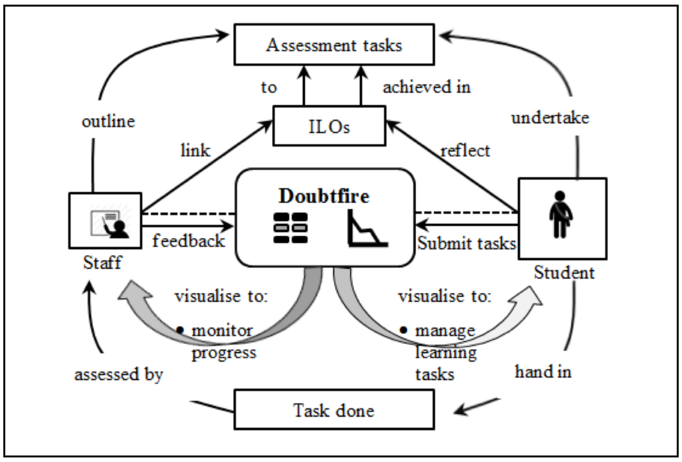

Fig. 1. Support for formative feedback in Doubtfire

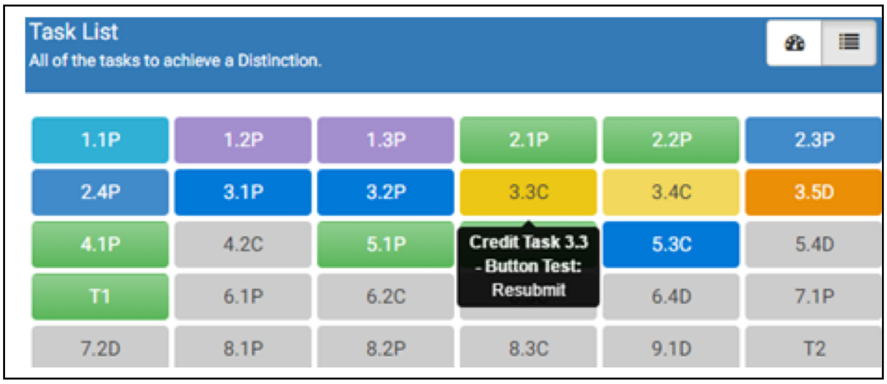

Fig. 2. Doubtfire task list

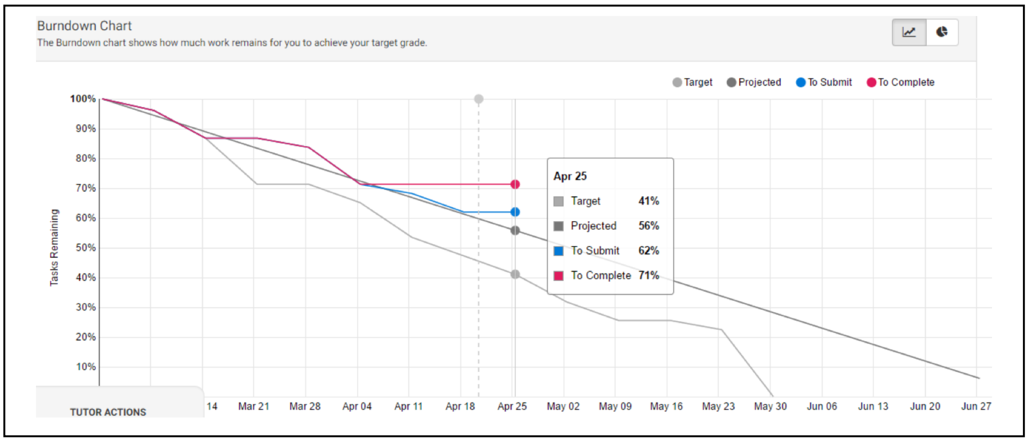

Fig. 3. Doubtfire burndown chart

With the establishment of the links through open learner model visualisations, Doubtfire++ could better support student reflections on task assessment in terms of learning outcomes. It encourages them to think about what they want to achieve or what they have achieved in attempting or completing an assessment task. It is also posited that the visualisations can help the teaching staff to gain useful insights of how students learn in this new environment. In this way, staff can provide appropriate guidance to students on time as well as help them to reflect for any changes needed to further improve their practice in a CA approach.With these enhanced features, the application as in Fig. 4, will support student reflections on task assessment and encourage them to think about what they have achieved in attempting and completing a unit tasks. In this, it is hoped that the tool will guide students to engage in high level cognitive activities and enhance their meta-cognitive skills. Doubtfire++ is believed to better support the frequent formative feedback in a Task-Oriented Portfolio teaching approach besides encouraging student self-regulated learning.

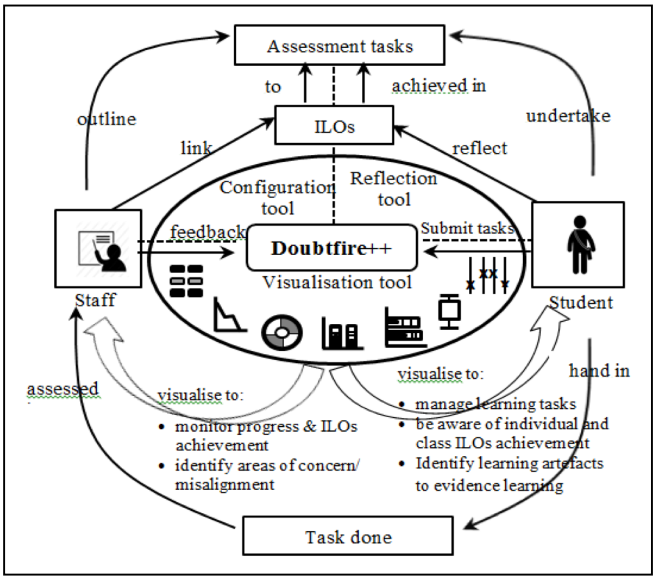

Fig. 4. Additional visualisations in Doubtfire++ supporting reflection on learning outcomes

A. Doubtfire++ Interface Design

With open learner model visualisations, Doubtfire++ supports staff with comprehensive learning analytics data about students in general, and provides insightful data about possible staff and student perceptions on the links between unit tasks and unit learning outcomes. These analysis and visualisations aim to support staff reflection on teaching by identifying how well teaching and learning strategies are working for students, including the teaching resources such as the content, materials and tools, and the assessment procedures. Some Doubtfire++ interface designs follow. Fig. 5 is an example of interface used by teaching staff to monitor class progress toward learning outcome achievements using box plots. The teaching staff can gain information about student current performance in terms of learning outcome achievements. At the end of the semester, the data helps the teaching staff to reflect if the learning activities provided have led to the required level of learning outcome achievements and if the achievements are in line with the unit requirement. This information can be used to guide staff future planning for the unit. Fig. 6 shows the interface through which students can inspect and compare the progress using an adapted version of Stephen Few's bullet chart. Four different colour hues are exploited to represent the qualitative range for different levels of achievement, indicating staff expectation for a Pass, Credit, Distinction and High Distinction. The graph also shows other quantitative attributes such as the class average, class achievement range, and an individual's achievement. Users are able to quickly grasp an overall understanding of an individual's achievement as well as the class achievement as compared to staff expectation. In this way, students are provided with a clear target to excel. Fig. 7 shows an example of the student reflection interface via which students align unit tasks with intended learning outcomes. Through this visualisation, students align unit tasks with the intended learning outcomes to demonstrate that they have achieved unit learning outcomes in preparing their portfolio for final assessment. It encourages students to reflect on their progress and achievements based on the tasks completed by clicking on the ratings that best represent their knowledge gained. They can then upload their work to showcase or evidence their achievements. While the visualisation facilitates student self-reflection, the reflection data captured can provide insight about student learning. This helps the teaching staff to examine any potential misalignment arising from mismatches between student reflections on learning and initial staff plans.

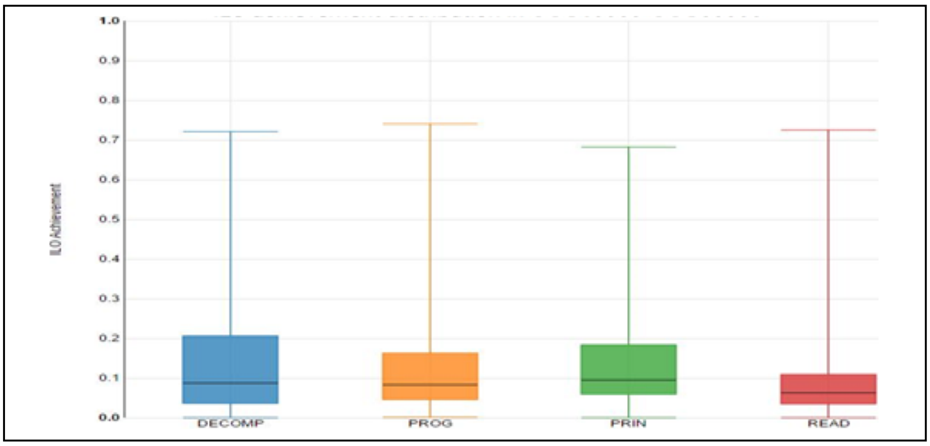

Fig. 5. Interface for teaching staff to monitor class LO achievements

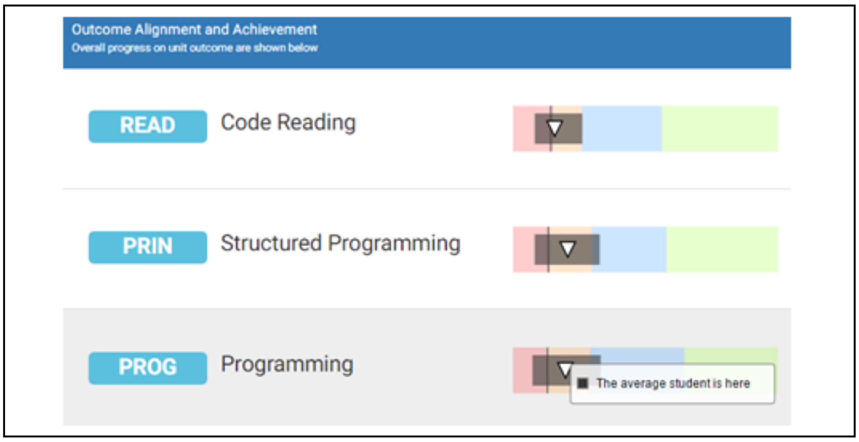

Fig. 6. Interface for a student to inspect and compare their progress toward achieving learning outcomes

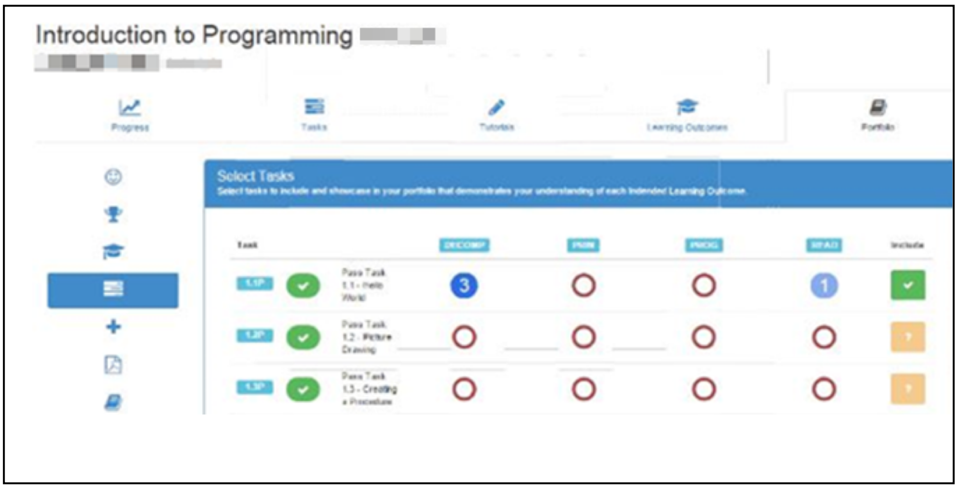

Fig. 7. Interface for a student to reflect their LO achievements based on the tasks completed

IV. Usability Testing

Usability of a software product can be measured by User eXperience (UX) metrics that reveal user experience interacting with a product [21]. UX metrics may include task success or failure, task completion time, errors made, effort required and learnability. This means UX metrics can reflect some aspects of the user personal experience using a product or system in a numeric format. These numerical values reveal the information about a user's interaction with the product. The aspect of effectiveness of a tool or product can be deduced from the situation where a user is able to use it to complete a task. The efficiency of a tool can be determined from the amount of effort spent – the number of clicks and time required in completing a task. User satisfaction can be known based on the degree a user is happy with his or her experience when performing a task with the product [21]. Therefore, understand usability issues from the UX metrics can reveal the performance of the tool and help to improve user experience through iterative design improvement and increase the buy-in of the product or tool in the future.

A usability study can be in the form of either formative or summative or both. Formative approach aims to collect data to improve the design of a product before it is launched or released whereas summative approach aims to find out to what extent the specified target goals have been achieved. According to Albert and Tullis [21], the earlier the formative evaluation is performed, the greater opportunity the usability evaluations can impact the design and usability positively.

A. *Usability Testing for Doubtfire++*

As Doubtfire++ is an enhanced version of Doubtfire, user interaction with it is an important aspect to be studied. This application aims to support student learning, and usability testing can help ensure its ease-of-use, and can be used effectively by both staff and students as a reflection tool. Through analyses and visualisations generated by Doubtfire++, students will be invited to reflect and regulate their learning whereas staff will be able to easily interact with the tool to gain insight about student learning and to reflect on their teaching. Therefore, we conducted a case study in the form of formative usability testing to evaluate Doubtfire while it is being designed and developed in order to identify any possible usability issues and to receive recommendations from prospective users to make iterative design improvements. This testing aimed to provide initial details to verify that these visualisations will support student learning, and assist staff in monitoring and supporting the learning process.

B. *Research Methodology*

This study aims to examine usability issues to improve user experience using Doubtfire++ through a formative evaluation based on UX metrics. Usability tasks were outlined based on Doubtfire improved features – those related to reflection and inspection of learning outcome achievements and linking of tasks to learning outcomes. UX metrics used in this study were focused in this aspect to quickly and easily collect the data as part of a normal iterative usability evaluation. Ten respondents were recruited from interested teaching staff and students in Swinburne University of Technology in Australia through self-selection sampling. There were 5 students, 4 tutors and a teaching staff in which 5 were males and 5 were females. All test participants had at least 5 years of experience using LMS.

C. Experimental Procedure

The formative usability testing involved a one-on-one session between a moderator and test participant to collect data through UX metrics and self-reported metrics. Before the participants started to perform the tasks, they were briefed about the objectives of this study and were given some time to read through all the tasks. They were aware of the teacher interface and the student interface. There was no training or walk-through of the tool. During the usability testing, each participant used the computer to interact with the prototype application. Their interactions with the tool were observed by the moderator. Data was collected in the form of UX metrics and participants' self-reported metrics. Screen captures were taken using Adobe Captivate 9 Software Simulation. For every task that the participants had completed, they were requested to say out loud their answers, which the moderator recorded in an Excel spreadsheet. The spreadsheet was formulated with error checking and time stamping functions to capture the metrics that included number of task success or failure, task completion time and participants' verbatim comments. Data regarding a participant's effort while interacting with the tool included the number of mouse clicks and screen navigation was collected through screen captures. After the participants had completed all the tasks, they were asked to complete a simple After-Scenario Questionnaire and to rate ease-of-task or satisfaction for each task in a paper questionnaire. This was followed by a Retrospective Think Aloud (RTA) session to discuss the task failures or errors. We sought their explanation in order to find out the possible cause and potential usability issues and to gain design recommendations to improve the prototype application.

D. Usability Tasks

There were 18 usability tasks distributed in 3 parts. The tasks were identified based on the improved features that aimed to gain insightful data about prospective user perceptions on the links between unit tasks and unit learning outcomes to encourage student-regulated learning. Part A consisted of teacher configuration tool to create tasks, outline learning outcomes and link tasks to these outcomes, denoted by A1 to A3. Part B listed 8 tasks that were labelled as B1 to B8. The tasks aimed to evaluate the shared teacher view interface to inspect and monitor student learning progress and LO achievements. Part C was a list of 7 tasks, labelled as C1 to C7 using the student view interface that included task for setting target grade, inspecting learning data, reflecting learning and uploading a portfolio for assessment. It is important to note that some of the tasks in teacher view interface in Part B are shared interface in which they can accessed by both teacher and students whereas tasks in Part C are solely meant for students.

E. User eXperience (UX) Metrics

UX metrics can reveal user experience interacting with the tool. Two common UX metrics are performance metrics and satisfaction metrics. Performance metrics refer to what a user does that include task completion rate, task time and the effort needed to complete a task. Task success rate gives insights about the effectiveness aspect of the tool. In this study, task success refers to the situation where a participant was able to complete a task without any help. The amount of effort used to complete a task was derived from the task completion time and number of clicks that also included the screen navigation counts when completing a task in order to determine the efficiency of the tool.

Some tasks are inherently more complicated than others. Thus the average task time and the average number of clicks or navigation count data may not provide meaningful reasoning for the possible usability issues. Instead, we established an acceptable threshold value for the time taken and effort needed to complete each task. Albert and Tullis [21] suggested double or triple the minimum amount of time or effort needed to complete a task to set the acceptable threshold value. The threshold values were set based on the time taken and the number of clicks performed by the moderator who knew how to complete the tasks. As the minimum time required to complete some tasks is relatively short, i.e. only 8 seconds is required to identify the LO achievement level and only 1 click is needed to view and interpret the burndown chart, any short delay or unintended click will have a great impact on the task time and the number of click data. Thus we tripled the time and the number of clicks when determining the threshold value for each task. The task completion time for each participant were then compared to this value. Any tasks that took more than the threshold were considered problematic.

F. Self-reported Metrics

Satisfaction concerns participant perceptions about their interaction with the prototype, including what they said, felt or thought with regards to their interaction with the tool. To investigate user satisfaction, participants' self-reported metrics were used to collect ease-of-task data for each task at the end of the usability testing. Participants were asked to rate each task with a five-point semantic differential scale ranging from very easy to very difficult. An After-Scenario Questionnaire, adopted from Lewis [22] with five-point Likert scale ranging from strongly disagree to strongly agree, was used to collect data about their overall experience interacting with the tool. The questionnaire contained only 3 questions that aimed to investigate fundamental usability aspects such that question 1 was related to effectiveness, question 2 was about efficiency aspect and question 3 was related to overall satisfaction that also included effectiveness, efficiency and participants' satisfaction on ease-of-task.

While the performance metrics – task success or completion time – could identify or uncover potential usability problems, they were supplemented with observational and self-reported data to explore the causes of the problems and the ways prototype can be improved.

V. RESULTS

This section presents the data analysis results from usability testing. As the sample size was relatively small, *n*=10, confidence intervals were used to reflect trust or confidence in the data. Data collected from UX metrics were binary data thus the Adjusted Wald values were used to denote the confidence interval at 95% level in the bar charts as shown from Fig. 8 to Fig. 10. Ease-of-task data collected from self-reported metrics are based on five-point Likert scale, thus error bars at 95% confidence level were used to indicate the confidence interval as shown in Fig. 11.

Fig. 8 shows successful completion rate by task for all 10 participants. The chart shows the comparison of the success rates for each task. Detailed analysis of each task was done by looking at the specific problems that caused task failure to determine the changes that may be needed to address the potential usability issues. There are 5 interfaces that 5 or more out of 10 participants failed to successfully complete the tasks. The interfaces are B1, B5, B6, C1 and C6. These interfaces appear to be problematic. Interfaces that work well include A1, A2, B3, B7 and C5. Other interfaces including A3, B2, B4, B8, C2, C3, C4 and C7 are likely to have usability issues. Participants' self-reported metrics were used to detect the potential problems.

Fig. 9 shows the percentage of participants that the task time exceeded the threshold value, by task. There are 4 tasks where more than 50% of the participants took an excessively long time to complete them, i. e., task B2, B5, B7 and C2.

Fig. 10 indicates the percentage of participants who needed an excessive number of clicks to complete each task. There are 7 tasks where more than 50% of the participants spent excessively more effort to complete them. The tasks are B2, B4, B5, B6, B8, C1 and C4.

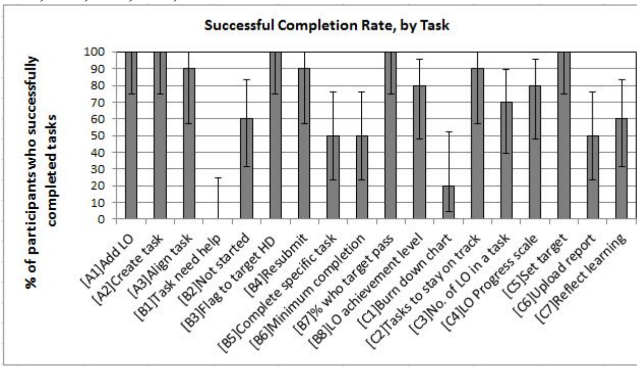

Fig. 8. Successful completion rate for each task

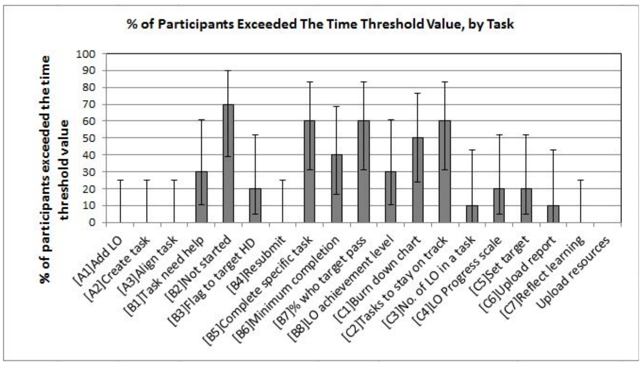

Fig. 9. Percentage of participants exceeded time threshold

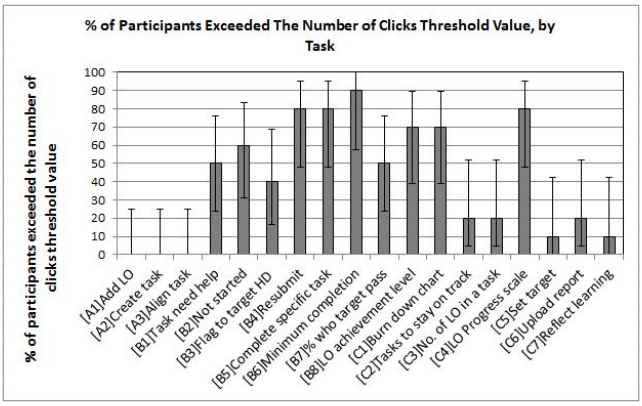

Fig. 10. Percentage of participants exceeded number of clicks

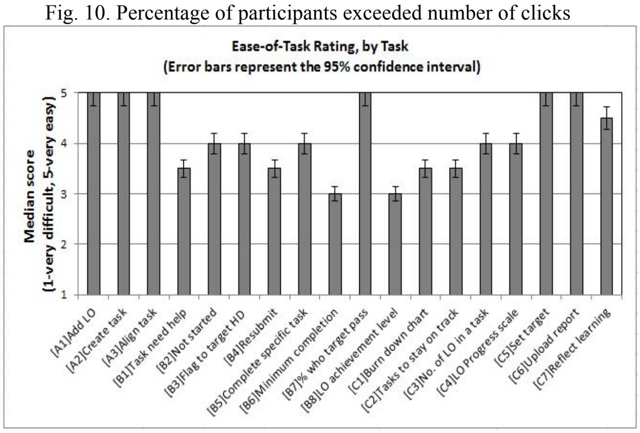

Fig. 11. Ease-of-task rating for each task

The ease-of-task ratings as in Fig. 11, are participants' self-reported metrics about their satisfaction regarding their interaction with the prototype tool. To avoid issues regarding outliers, the median score is used to present the data. Six tasks that have median score less than 4 are B1, B4, B6, B8, C1 and C2.

Table I shows a summary of participants' responses, mean and median scores regarding their overall experience interacting with the tool collected from After-Scenario Questionnaire. Based on the data obtained from the performance metrics and satisfaction metrics, we further analysed the data to investigate the correlations between the three measures of performance, i.e. the task success, task time and number of clicks and one measure of satisfaction, the ease-of-task rating. The results are as shown in Table II. We also hypothesized that there is a positive correlation between the 2 measures of performance, the task time and number of clicks. Table III shows the result.

TABLE I. PARTICIPANTS' RESPONSES ABOUT THEIR OVERALL EXPERIENCE

Item	Very Easy		Very Difficult			Mean, µ	Median
	(5)	(4)	(3)	(2)	(1)		
Effectiveness	2	3	4	1	0	3.6	3.6
Efficiency	2	3	4	1	0	3.6	3.6
Satisfaction	3	3	2	2	0	3.7	4.0

TABLE II. CORRELATIONS BETWEEN PERFORMANCE AND SATISFACTION MEASURES

	1	2	3	4
1. Task success	–	–0.294	–0.391	0.504*
2. Exceed time thres.		–	0.517*	–0.531*

3. Exceed click thres.			–	–0.794**
4. Ease-of-task				–

* Correlation is significant at the 0.05 level (2-tailed)
** Correlation is significant at the 0.01 level (2-tailed)

TABLE III. CORRELATIONS BETWEEN TIME NEEDED TO COMPLETE A TASK AND THE NUMBER OF CLICKS PERFORMED

	Between time and the number of clicks
Pearson Correlation	.756**
Sig. (2-tailed)	.000
N	18

** Correlation is significant at the 0.01 level (2-tailed)

VI. DISCUSSION

Data from Figure 8 to Figure 11 are summarized in a matrix form in Fig. 12. Interfaces with potential usability issues identified from data presented from Fig. 8 to Fig. 11 were marked in this matrix. This matrix thus gives an overall picture about problematic interfaces that need further enhancement and improvement. None of the task has 4 checked marks. Usability tasks that have 3 checked marks are highlighted in red, yellow denotes having 2 checked marks whereas brown for 1 checked mark. Tasks without any checked marks indicate interfaces that work well and participants were able to successfully complete the tasks easily at a reasonable time and they were satisfied with the interfaces. Participants' verbatim comments during the usability testing and during the RTA session and the moderator's observation comments were analysed to gain insights about the source of the problem and to find out the aspects of the tool that need improvement.

A. Potential Usability Issues Based on Usability Metrics

1) Usability Issues in Shared Teacher Interface

As can be seen from Fig. 12, potential usability issues are clustered in the shared teacher view interfaces (B1 – B8). In B1 interface, there are 3 icons to view task status in the Feedback tab. Each icon consists of data about students in a particular lecturer's class or all students enrolled in a unit. By default, only students who have registered in a particular class will be shown. To view the data for all students, an icon "all students" will have to be clicked. Most participants were stuck in this step as they were no aware of the "all students" icon. The interface in B2 has sorting function to sort students' task status. However, most participants were unaware of this sorting feature and completed the task by navigating through all pages in the list. Thus, this task recorded high rate of exceeding time and/or click threshold. To improve user experience when interacting with the prototype tool, participants suggested having a short training or demonstrate session to show the features available that can provide specific data or can ease the inspection of learning data such as the show "all students" icon and the sorting function.

B5 to B8 tasks are about the Analytics tab that provides class statistics data including task status distribution, task completion, student target grade, learning outcome achievement and task statistics. These tasks required participants to have analytical mind and basic knowledge in statistics to interpret data presented graphically. The matrix in Fig. 12 shows high percentage of participants exceeded the time and/or click threshold when completing these tasks. The main reason may be due to the hover over tool tips in each icon. As there are 5 icons in this tab, participants hovered over around the tool tips to see the description of the icons. Through observation and participants' verbatim comments, not many participants were able to interpret the box plot. Participants suggested labelling the icons instead of providing hover over tool tips. We also plan to have a part that can explain how to interpret box plot or include this explanation in the training when we introduce this tool to the new users.

2) Usability Issues in Student Interface

Tasks listed from C1 to C7 are interfaces in the student view. The major issue found was interpreting burndown chart (Fig. 3) that showed the number of tasks completed by week or remaining by week. Only 2 out of 10 participants were able to complete this task without any help with correct interpretation. By default, there are 4 line graphs in this chart showing target and projected completion time, percentage of tasks submitted and completed. A user can toggle to turn the line on to be visible on the chart or turn it off. The burndown chart decreases as work has been completed. This means when 10% of the tasks have been completed, the "complete" line will show 90%. This may have caused misinterpretation that 90% have been completed. We will either modify the wording or provide a statement to explain the "complete" line in the interface. The other way to amend this could be by labelling the y-axis as 0% from the top and ends at the bottom in 100%.

Another issue is about task failure in C2. Nine out of 10 participants did not notice the tasks to be done, as indicated just beside the burndown chart. This information is useful for student who may be lagged behind to focus on a certain tasks in order to get back on track. We will need to make it more noticeable or specifically highlight this feature to students during introduction session or training. Besides that, the hover over tool tip problem was also encountered by participants when interpreting the LO progress scale. This interface recorded high percentage of participants exceeded the click threshold. Participants had to hover over the scale to see the graphical representation such as colour, line and triangle. To improve the interface design, 3 participants suggested having legends to describe the graphical representation. In addition, as the upload report function in task C6 only accept drag and drop function to upload report, participants had to drop the file in the drop box first before they could upload it. Nevertheless, due to the drop box is in dull grey colour as compared to the bright green "Upload" button that is more noticeable, participants tended to click the "Upload" button instead of dragging the file into the drop box. Some kept on clicking the "Upload" button and could not get the file uploaded. There was no error message informing a user to drop the file before it can

Usability tasks / Usability metrics	A1	A2	A3	B1	B2	B3	B4	B5	B6	B7	B8	C1	C2	C3	C4	C5	C6	C7
High task failure rate				x				x	x			x					x	
Exceed time threshold					x			x		x			x					
Exceed click threshold					x		x	x	x		x	x			x			
Ease-of-task rating < 4				x			x		x		x	x	x					

Fig.12. Matrix of usability tasks and the usability metrics

be uploaded. Verbatim comments received to improve the tool included having the "upload" button to be linked to "browse folder to upload file" or remove the upload button and to just show it after a user has dropped the file, or having a pop up message to tell users that they need to "Drop Report" first when the "Upload" button is clicked.

B. Potential Usability Issues Based on Moderator's Observation

Apart from the issues found based on the usability metrics, there are a few issues found based on the moderator's observation. The calendar and the delete task icons that are at the far end of the date input box and far apart down the tasks created respectively make them imperceptible. Not all participants noticed these functional icons and use them during the usability test. We will position the icons appropriately. Besides that, one bug issue has been observed and will be rectified.

C. User Experience

Participants' overall experience interacting with the tool was quite positive and encouraging. The mean and median scores for effectiveness and efficiency are both 3.6 out of the scale of 5 respectively. For satisfaction, the mean and median scores are 3.7 and 4.0 respectively. Although participants encountered problems with some of the interfaces when completing the usability tasks, they were quite satisfied with the effectiveness and efficiency of the tool and were also quite satisfied with the ease-of-use of the tool.

D. The Relationships between Performance and Satisfaction Metrics

We investigated the relationships of performance metrics that include the task success rate, percentage of participants exceed time threshold, percentage of participants exceed click threshold with the satisfaction metrics, the ease-of-task ratings, by using Pearson product-moment correlation coefficient. Although Cohen [23] suggested the $0.30 < r < 0.49$ as medium strength of correlation, only correlations with large strength were found to be statistical significance, i.e. $r > 0.50$, due to relatively small number of the usability tasks, $n=18$ for moderate correlations to reach statistical significance [24].

We found that there was a strong, positive correlation between task success rate and ease-of-task rating, $r=0.504$, $n=18$, $p<0.025$ with more participants successfully completed the tasks with higher satisfaction. Besides that, very strong, negative correlation was observed between percentage of participants exceed click threshold and ease-of-task rating, $r=-0.794$, $n=18$, $p<0.005$, with high percentage of participants exceed click threshold with lower satisfaction. In addition, there was also a strong, negative correlation between percentage of participants exceed time threshold and ease-of-task rating, $r=-0.531$, $n=18$, $p<0.025$, with high percentage of participants exceed time threshold with lower satisfaction. These results suggest a relationship exists between performance and satisfaction measures. This is useful finding. It indicates that user satisfaction is closely related to the efficiency and effectiveness of the tool. Enhancing the prototype tool to minimize the amount of time and effort required in completing a task is essential to improve user experience using the tool.

We also observed very strong, positive correlation between the time required to complete a task and number of clicks performed, $r=0.756$, $n=18$, $p<0.005$ with more time spent in completing a task with more number of clicks performed. This also can be seen from the strong, positive correlation between the percentage of participants exceed time threshold and percentage of participants exceed number of clicks threshold values, $r=0.517$, $n=18$, $p<0.005$ with more excessive time spent in completing a task with more excessive number of clicks performed. This indicates the importance of ease-of-use interface. If the interface is easy to use, users will be able to complete their tasks with minimum effort and time. This will improve user satisfaction and experience.

E. Threats to Validity

The test participants were recruited from self-selection sampling. They could be users who liked this study or the prototype tool hence leading to exaggeration of the research findings. Besides that, some tasks required analytical mind and basic statistical knowledge to interpret the graphical data. This could have introduced bias to the results obtained.

F. Summary findings

This usability case study in the context of teaching philosophy has opened up a different set of considerations. The results of this case study allow us to tackle the potential usability issues and improve the application so that the actual target users will have a better experience interacting with the tool when it is deployed in any teaching and learning process. The key learnings from this case study include:

- Key functional icons for a system have to be made noticeable to users. This can be done through highlighting the icon or use of a blinking icon or through a demonstration.
- For a system that involves data analytics or presentation of graphical representations, it is essential to always have a legend or an example to explain the encoded data.
- Proper use of colors and pop out messages to guide users for further actions and lead to a better human computer interaction experience.

- Static text is an alternative way, sometimes even a better way, to indicate the function of a functional icon than a hover over tool tip in a system.
- A relationship exists between performance and satisfaction measures. If an interface is easy to use, tasks can be completed with a minimum effort and time. This will improve user satisfaction and their interaction experience.

VII. Conclusion

We see the potential of OLM visualisations in supporting constructive alignment to improve learning outcomes. We realised the need to have a software tool addressing learning outcome to improve learning. We have enhanced Doubtfire to be a more robust visualisation tool with new visualisations to help staff and students monitor progress toward achievement of learning outcomes. User experience interacting with the prototype is a critical factor in determining the success or failure of a new product or tool. We conducted usability testing via the use of usability metrics and self-reported metrics to detect potential usability issues in order to tackle any problems and to determine if these changes are likely to support both the teaching staff and student in the teaching and learning process. We have gained valuable insights about potential usability issues and recommendations to improve the application. We now have a better understanding about the performance of the tool. This includes the interfaces that work well and do not as well as parts that need to be emphasized during an introductory training sessions, how to make the functional icons to be noticeable to users . Our next step is to improve and deploy the updated application to investigate support for an outcome-based student-centred teaching and learning approach.

Acknowledgment

We gratefully acknowledge Multimedia University and Swinburne University of Technology for support for this research. We thank all of the respondents and participants.

References

[1] M. Richter and M. Flückiger, *User-Centred Engineering: Creating Products for Humans*: Springer, 2014.

[2] ISO/DIS, "9241-210: 2010. Ergonomics of human system interaction-Part 210: Human-centred design for interactive systems," *International Standardization Organization (ISO). Switzerland,* 2010.

[3] F. Mahnegar, "Learning management system," *International Journal of Business and Social Science,* vol. 3, pp. 144-150, Jun. 2012.

[4] L. A. Wankel and P. Blessinger, "New vistas in higher education: An introduction to using social technologies," in *Increasing Student Engagement and Retention Using Social Technologies: Facebook, E-portfolios and Other Social Networking Services (Cutting-edge Technologies in Higher Education)*, L. A. Wankel and P. Blessinger, Eds., ed Bingley, U.K: Emerald Group Publishing, 2012, pp. 3-16.

[5] S. F. Tello and L. Motiwalla, "Using a learning management system to facilitate learning outcomes assessment," in *Learning Management System Technologies and Software Solutions for Online Teaching: Tools and Applications*, Y. Kats, Ed., ed Hershey, PA: Information Science Reference, 2010, pp. 138-156.

[6] A. Cain, "Constructive alignment for introductory programming," Faculty of Information and Communication Technologies, Swinburne University of Tecnology, Australia, 2013.

[7] J. Biggs, "Constructive alignment in university teaching," *HERDSA Review of Higher Education,* vol. 1, pp. 5-22, 2014.

[8] C. J. Woodward, A. Cain, S. Pace, A. Jones, and J. F. Kupper, "Helping students track learning progress using burn down charts," presented at the 2013 IEEE International Conference on Teaching, Assessment and Learning for Engineering, Bali, Indonesia, 2013.

[9] J. A. Self, "Bypassing the intractable problem of student modelling," in *Intelligent Tutoring Systems: At the Crossroads of Artificial Intelligence and Education*, C. Frasson and G. Gauthier, Eds., ed Norwood, N.J: Ablex, 1990, pp. 107-123.

[10] S. Mohanarajah, R. Kemp, and E. Kemp, "Opening a fuzzy learner model," in *12th International Conference on Artificial Intelligence in Education: Workshop on Learner Modelling for Reflection, to Support Learner Control* Amsterdam, The Netherlands, 2005, pp. 62-71.

[11] D. Hartley and A. Mitrovic, "Supporting learning by opening the student model," in *Intelligent Tutoring Systems*, 2002, pp. 453-462.

[12] S. Bull, V. Dimitrova, and G. McCalla, "Open learner models: Research questions," *International Journal of Artificial Intelligence in Education,* vol. 17, pp. 83-87, 2007.

[13] J. Kay, Z. Halin, T. Ottomann, and Z. Razak, "Learner know thyself: Student models to give learner control and responsibility," in *International Conference on Computers in Education*, Kuching, Malaysia, 1997, pp. 17-24.

[14] S. Bull and J. Kay, "Student models that invite the learner in: The SMILI open learner modelling framework," *International Journal of Artificial Intelligence in Education,* vol. 17, pp. 89-120, 2007.

[15] S. Bull, T. J. Jackson, and M. J. Lancaster, "Students' interest in their misconceptions in first-year electrical circuits and mathematics courses," *International Journal of Electrical Engineering Education,* vol. 47, pp. 307-318, 2010.

[16] S. Bull and J. Kay, "Open learner models," in *Advances in Intelligent Tutoring Systems*, R. Nkambou, J. Bourdeau, and R. Mizoguchi, Eds., ed Berlin Heidelberg: Springer-Verlag 2010, pp. 301-322.

[17] Y. W. Sek, H. Deng, E. McKay, and M. Qian, "Investigating the impact of learners' learning styles on the acceptance of open learner models for information sharing," in *Australasian Conference on Information Systems*, Adelaide, Australia, 2015.

[18] S. Girard, "Towards promoting meta-cognition using emotive interface personas within Open-Learner Modelling Environments," in *Proceedings of the 2007 conference on Artificial Intelligence in Education: Building Technology Rich Learning Contexts That Work*, 2007, pp. 687-688.

[19] C. Y. Law, J. Grundy, A. Cain, and R. Vasa, "A preliminary study of open learner model representation formats to support formative assessment," in *IEEE 39th Annual International Computers, Software & Applications Conference*, Taiwan, 2015, pp. 887-892.

[20] C. Y. Law, J. Grundy, A. Cain, and R. Vasa, "An empirical study of user perceived usefulness and preference of open learner model visualisations," presented at the IEEE Symposium on Visual Languages and Human-Centric Computing, Cambridge, UK, 2016.

[21] W. Albert and T. Tullis, *Measuring the user experience: Collecting, analyzing, and presenting usability metrics*, Second Edition ed. United States of America: Elsevier Inc., 2013.

[22] J. R. Lewis, "Psychometric evaluation of an after-scenario questionnaire for computer usability studies: the ASQ," *ACM SIGCHI Bulletin,* vol. 23, pp. 78-81, 1991.

[23] J. Cohen, "Statistical Power Analysis for the Behavioral Sciences. 2nd edn. Hillsdale, New Jersey: L," ed: Erlbaum, 1988.

[24] J. Pallant, *SPSS survival manual*: McGraw-Hill Education (UK), 2013.