Computers and Education: Artificial Intelligence 3 (2022) 100069

Contents lists available at ScienceDirect

Computers and Education: Artificial Intelligence

journal homepage: www.sciencedirect.com/journal/computers-and-education-artificial-intelligence

Enhancing learning by Open Learner Model (OLM) driven data design

Judy Kay [a,*], Kathryn Bartimote [b], Kirsty Kitto [c], Bob Kummerfeld [a], Danny Liu [d], Peter Reimann [b]

[a] School of Computer Science, The University of Sydney, Australia
[b] School of Education and Social Work, The University of Sydney, Australia
[c] Connected Intelligence Centre, University of Technology Sydney, Australia
[d] The University of Sydney, Australia

ARTICLE INFO

Keywords:
Open Learner Models (OLMs)
Learner models
Student models
Scrutability
Ontologies
Self-regulated learning
Learning analytics

ABSTRACT

There is a huge and growing amount of data that is already captured in the many, diverse digital tools that support learning. Additionally, learning data is often inaccessible to teachers or served in a manner that fails to support or inform their teaching and design practice. We need systematic, learner-centred ways for teachers to design learning data that supports them. Drawing on decades of Artificial Intelligence in Education (AIED) research, we show how to make use of important AIED concepts: (1) *learner models*; (2) *Open Learner Models (OLMs)*; (3) *scrutability* and (4) *Ontologies*. We show how these concepts can be used in the design of OLMs, interfaces that enable a learner to see and interact with an externalised representation of their learning progress. We extend this important work by demonstrating how OLMs can also drive a learner-centred design process of learning data. We draw on the work of Biggs on constructive alignment (Biggs, 1996, 1999, 2011), which has been so influential in education. Like Biggs, we propose a way for teachers to design the learning data in their subjects and we illustrate the approach with case studies. We show how teachers can use this approach today, essentially integrating the design of learning data along with the learning design for their subjects. We outline a research agenda for designing the collection of richer learning data. There are three core contributions of this paper. First, we present the terms OLM, learner model, scrutability and ontologies, as thinking tools for systematic design of learning data. Second, we show how to integrate this into the design and refinement of a subject. Finally, we present a research agenda for making this process both easier and more powerful.

1. Introduction

Open Learner Models (OLMs) are a core idea that has emerged from AIED research (Bull, 2020; Bull & Kay, 2016). From a student's perspective, an OLM provides a useful view of their learning progress as they master key learning objectives. From the perspective of the designer of AIED systems, the OLM is an interface that presents a view of the *learner model*. The learner model is often described as the AIED system's "beliefs" about the learner's knowledge, misconceptions, preferences, goals and attributes. In a classic AIED system, the main purpose of the learner model is to drive personalisation of the learning interface. In this paper, we take these ideas from AIED, and we show how teachers of a typical university subject can use them in a systematic design of the learning data that can enable students to track their learning progress.

Our approach provides a path towards making the design of learning data a standard part of the learning design process for a subject. This is a major shift from current thinking about learning data, as reflected in the literature, which is to begin with the data that is readily available and then analyse it to gain insights that can enhance learning. Current educational technology certainly can generate large volumes of learning data about each student. Indeed, recognition of the potential value of such data has spurred the fast growth of the learning analytics research community. While that community has been able to extract valuable insights from learning data, there has been relatively little work on learner facing dashboards (Bodily et al., 2018). It is now timely to create ways for every teacher to integrate the design of the learning data for their subject in a form that can help students appreciate the learning objectives set for the subject and to track their learning progress in terms of these.

Essentially, we present a new approach for teachers to use when they design their university subject. This starts with the teacher designing the

* Corresponding author.
E-mail addresses: Judy.Kay@sydney.edu.au (J. Kay), kathryn.bartimote@sydney.edu.au (K. Bartimote), kirsty.kitto@uts.edu.au (K. Kitto), bob.kummerfeld@sydney.edu.au (B. Kummerfeld), danny.liu@sydney.edu.au (D. Liu), peter.reimann@sydney.edu.au (P. Reimann).

https://doi.org/10.1016/j.caeai.2022.100069
Received 30 July 2021; Received in revised form 7 April 2022; Accepted 7 April 2022
Available online 2 May 2022

OLM they want their students to see. Then they design the data needed for the OLM design, integrating this with the established and widely used approach of constructive alignment (Biggs, 1996, 1999). Over decades, constructive alignment has been used by teachers to align the learning objectives of a subject with its teaching and learning (T&L) activities and assessments. Thus constructive alignment has two roles. First, it ensures that the learning design of a subject provides students with adequate opportunities to develop the intended knowledge, skills and attitudes. Secondly, it drives the design of the assessments so that they provide both formative and summative feedback about each of them. With the growing role of learning data in education, this paper shows how to integrate the design of learning data with this overall learning design process.

We aim to make three key contributions. First, introduce a new way to make use of the AIED terms, OLM, learner model and scrutability, as a part of designing a subject and its learning data. Second, we illustrate how this approach can be used with current technology to harness data from widely used learning technology, such as Learning Management Systems (LMSs), videos, self-test resources and automated grading software. This has a theoretically grounded discussion of the ways that teachers can ensure that an OLM created by our approach is integrated into, and benefits, the learner. The third contribution is a research agenda that describes the tools that would make it quick and easy for teachers make rich uses of our approach.

The next section introduces background research that informed our work. Then we provide a high level overview of our approach, followed by case studies of the use of the approach in two stages: a data design process that is learner-centred, adheres to constructive alignment principles, and supports self-regulated learning; an example of one way to implement the detailed data collection and management as well as the presentation of the OLM. We then discuss the approach and research directions in terms of three key stakeholders: students, teachers and researchers, with educational theories, contexts and uses and our final section presents a summary and conclusions.

2. Background

This section presents the previous research that underpins this work. The first three parts are what we argue should be the educational foundations of OLMs. These begin with research on learning design that is learner centred and the importance of self-regulated learning in tertiary education. Then, we introduce Biggs' constructive alignment (Biggs, 1996, 1999) and the way it fits into that body of work. The next part introduces key ideas from AIED research on Learner Models and Open Learner Models (OLMs) and the notion of scrutability.

2.1. Learner-centred learning design

Learning design is defined by Bakharia et al. (2016) as a field that: "allows educators and educational researchers to articulate how educational contexts, learning tasks, assessment tasks and educational resources are designed to promote effective interactions between teachers and students, and students and students, to support learning …" Being learner-centred refers to qualities of the design as an outcome as well as to designing as a process. For the latter, being learner-centred means to include students in the design process, such as when university students co-design a subject with the teacher (Deeley & Bovill, 2015). It can also involve adaptation during the teaching (Reigeluth, Myers, & Lee, 2016). For example, a formative assessment in Week 3 can inform a student's choice between two learning activities in Week 4. A design is learner-centred if it enables educators to regularly answer the question: "Is an adjustment needed, and if so, what should that adjustment be?" (Popham, 2008). For Tertiary Education, at least, we may want to add a third element to our definition of learner-centredness: That this information also is made available to the students so that they can make informed adjustment decisions themselves (Kitto, Lupton, Davis, & Waters, 2017).

A student needs a criterion to answer the first part of the question, i.e. to assess whether an adjustment is needed. There is general agreement that the key criterion is attainment: what has been learned so far (Reigeluth et al., 2016). Many have argued – amongst them prominently Ference Marton (e.g., Marton, Hounsell, and Entwistle (1997)) and John Biggs (e.g., Biggs (1999)) – that attainment criteria should not only refer to content, but also to the ways students engage with the content, such as is captured in Bloom's Taxonomy (Anderson, Bloom and others, 2001) and the SOLO taxonomy (Biggs & Collis, 2014).

If a teacher or student determines that an adjustment is needed, the second part of the question requires that the learning design takes account of adaptations of instructional parameters along a number of dimensions. A very important dimension is learning time; in learner-centred instruction, learning time is individualised and made dependent on attainment, and level of mastery (Bloom, 1974). Other important dimensions are sequence of topics and learning/teaching method; to the extent possible, they too should be made contingent on attainment and learner preferences (Corno, 2008; Park & Lee, 2004). But why the interest in learner-centred design? One of the underlying motivations for this work is that it could help to encourage self regulated learning.

2.2. Self-regulated learning in tertiary education

Theoretical models of self-regulated learning (SRL) emerged in the 1990s seeking to bring together an explanation of learning regulation that included cognitive processes alongside those of motivation and emotion. We take the definition here that SRL is "an active, constructive process whereby learners set goals for their learning and then attempt to monitor, regulate, and control their cognition, motivation, and behaviour, guided and constrained by their goals and the contextual features in the environment" (Pintrich (2000):453). This gels well with the ethos of learner-centredness outlined in the previous section. Further, this definition of SRL relates to the self-regulatory aspect of human agency outlined in social cognitive theory (Bandura, 2005), in which a person's learning (and other) behaviour is the result of interactions between personal factors and the learning environment. The learning environment here includes the subject as designed by the teacher, and also the OLM used to provide feedback to students on their learning.

Panadero (2017) provides an overview of important SRL models. They differ in their relative emphasis on metacognitive versus motivational processes, and in the degree to which they articulate the interaction between the two. The most influential model is Zimmerman's (Zimmerman & Moylan, 2009). It depicts SRL as a goal-driven process comprising three cyclical and separate phases, each phase with clearly articulated sub-components. Its description is more at the macro-than micro-process level of description; and this helps when thinking about teaching and data design at the whole-of-subject level, which is the focus of this paper. See also (Azevedo, 2020) for an overview of current issues in the field.

The three phases in Zimmerman's model are forethought, performance, and self-reflection. In the forethought phase, students are engaged in task analysis activities of goal setting and strategic planning, and these are influenced by motivation beliefs of self-efficacy, outcome expectations, interest, value, and goal orientation. In the performance phase, students exert self-control in a range of ways including particular task strategies, self-instruction, imagery, time management, environmental structuring, help-seeking, interest incentives, and self-consequences. Also in the performance phase, students undertake self-observation via meta-cognitive monitoring and self-recording. In the self-reflection phase, one experiences a range of self-reactions including emotions, degree of self-satisfaction, and may take an adaptive or defensive stance as a result of learning experiences and outcomes. Also in the self-reflection phase, self-judgment includes self-evaluation and causal attribution for outcomes.

OLMs have the potential to support SRL, depending not only on the information shown, but also in how the OLM is used by students and teachers as an intermediary tool for communicating about goals, expectations, progress, achievement, etc. One of the defining properties of OLMs is that they give students external information that can used as feedback on satisfactory completion of tasks and/or performance. For example, in the earliest skill-meters Corbett and Anderson (1994) showed students their changing score on a small set of skills within a sophisticated intelligent tutoring system. They can act as an aide to meta-cognitive monitoring, but also as a prompt to enact particular self-control strategies such as help-seeking or (re)structuring one's environment. Beyond this, there have been systems that enable the student to contribute their own ratings regarding their mastery of a particular concept or objective (Cook & Kay, 1994). Other work enables the student to challenge a system rating and then the system provides them with the opportunity to complete tasks that demonstrate mastery (Bull, Brna, & Pain, 1995). There has also work where students set the standard for their learner model mastery (Upton & Kay, 2009). But how might we feed this work on learner centred and self-regulated learning into the rich history of OLMs? Here, we will argue that the bridge between the two fields can be provided by constructive alignment.

2.3. Constructive alignment

Biggs' notion of constructive alignment (Biggs, 1996, 1999, 2011) has been influential in curriculum design for tertiary education. This is a simple, elegant idea that is a powerful conceptual tool for teachers. Broadly similar approaches have been proposed by others. Notably Anderson (2002) reviews work on the similar notion called curricular alignment and instructional alignment is the term used in Cohen (1987) which reported large learning gains when it is well done. These bodies of work have different underlying educational theories. Our work particularly draws on Biggs' description of the ways that teachers can tackle curriculum design.

Fig. 1 gives a high level view of a subject that is constructively aligned. Essentially, constructive alignment calls for teachers to begin the design of a subject by defining the learning objectives, shown in the middle of the figure. The left of the figure shows the teaching and learning (T&L) activities. In constructive alignment, the teacher designs each T&L activity so that it can enable students to achieve one or more of the learning objectives. It is becoming common practice in universities for teachers to capture this reasoning in curriculum planning tools like Akari, CourseLoop or U-Planner,[1] as these enable a teacher to explicitly state the mapping between activities and learning objectives. For example, a teacher will typically describe each lecture and link it to one of more of the learning objectives. There is a huge diversity of possible activities designed by the teacher, such as the widely used lectures, tutorials and labs of tertiary education.

T & L activities increasingly use a wide variety of digital tools. The figure shows examples of common ones. But there are many others. Some were explicitly designed for teaching and learning (e.g. electronic textbooks). But many general purpose tools are also used in teaching – the figure shows the examples of videos and web pages. Each of these has the potential to provide learning data, and we will return to this point in Section 3.

In constructive alignment, the teacher must align the learning objectives with their design of the assessments (at the right of Fig. 1) and feedback. In the case of formative assessments, such as low stakes quizzes and class activities, this ensures that the students have feedback on their learning progress. With suitable learning design, students and teachers can be alerted to the need for remedial action. At the other end of the performance spectrum, students who are progressing well should gain reassurance that their current approaches to study are on track. Summative assessments, such as final exams, should serve both to provide grades and to enable the teacher to determine the success of their teaching strategy for each of the main learning objectives.

To this point, we have mainly taken a teacher-centred view, with a focus on the learning design. We now turn to a learner-centred perspective. The figure shows the learner at the top of the figure. The lines indicate that the learner should be able to see the learning objectives. Importantly, they should be able to see how the T&L activities and the assessments align with them. Modern curriculum design systems gives students this information. In this paper, we introduce OLMs as a different, personalised and dynamic visualisation of a student's progress against the teacher's learning objectives, potentially alongside information on progress towards their own additional learning goals.

2.4. Learner models and Open Learner Models (OLMs)

Open learner models have grown out of decades of AIED research. Initially, this was inspired by a vision of achieving the 2-sigma learning gains of expert 1:1 human tutors compared with classroom teaching (Bloom, 1984). Central to that vision was the learner model, which has been defined as *a machine's set of "beliefs" about the learner* (Wahlster & Kobsa, 1989). It is the learner model that drives the personalisation of an intelligent tutoring system.

2.4.1. A brief introduction to OLMs

Within even the early AIED work, it became common to provide the learner with an interface showing their progress, as a skill-meter (Corbett & Anderson, 1994). The broad idea of making the learner model available to the learner was proposed by John Self (Self, 1990) who later described 15 forms of openness (Self, 1999) in AIED systems. Such "opening" of the learner model became a strong theme in AIED research (Bull & Kay, 2016). Initially, most learner models and OLMs were part of personalised teaching systems. But *independent learner models* (Cook, Kay, & Kummerfeld, 2015; Kay, 1995) were also created, both for reuse of the learner model across personalised teaching systems and to support learners in self-monitoring, reflection and planning (Bull & Kay, 2013). We build on that work in this paper, taking the core ideas of a learner model and OLM as drivers for the systematic design of learning data in conventional university subjects that make use of typical teaching and assessment methods.

We now introduce OLMs by way of an example. Fig. 2 shows an OLM similar to that in (Brusilovsky, Somyürek, Guerra, Hosseini, & Zadorozhny, 2015). It has been used by many student cohorts in voluntary practice activities in a programming subject. The green cells across the top show the learner's progress in the topics listed above them. The darker the green, the higher the mastery level. So, for example, in the figure the *Variables* cell is dark green, indicating the student has mastered it.

This OLM uses *social comparison* – the blue cells show the learning progress of a "Group". (This introduces all the risks that comparison entails (Hanus & Fox, 2015; Khan & Pardo, 2016).) The middle row of the OLM makes it easier to compare this learner's performance against the group. Overall, the student has made similar progress to the group (indicated by the white cell for OVERALL). But on *Variables*, the green indicates they are doing better. The rest of the learning topics have a mix of cases where the student is ahead (green cells), similar (grey cells) or behind the "Group" (blue cells).

The "Group" could be the whole class, enabling the student to see themself compared with the class, or the "group" could be just a part of the class. For example, if a student is aiming for a high grade, they may want to see the group of high achieving students. The same visualisation could present a different standard for comparison. For example, the teacher could define the standard that is expected at this time in the semester, as in work where the teacher defined a "plausibly ideal student" (Cook et al., 2015; Kay & Lum, 2005). There are many other ways to help students make sense of their performance. For example, in a

[1] https://www.akarisoftware.com/https://uplanner.com/en/or https://courseloop.com/.

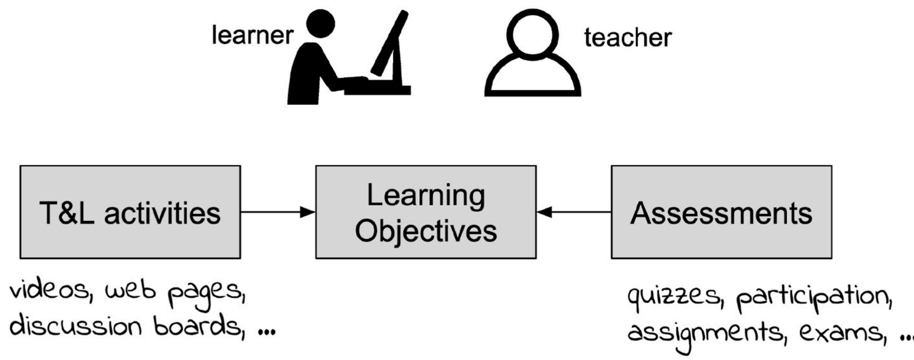

Fig. 1. Overview of constructive alignment. The boxes are the main parts that a teacher creates. All of these are visible to the student.

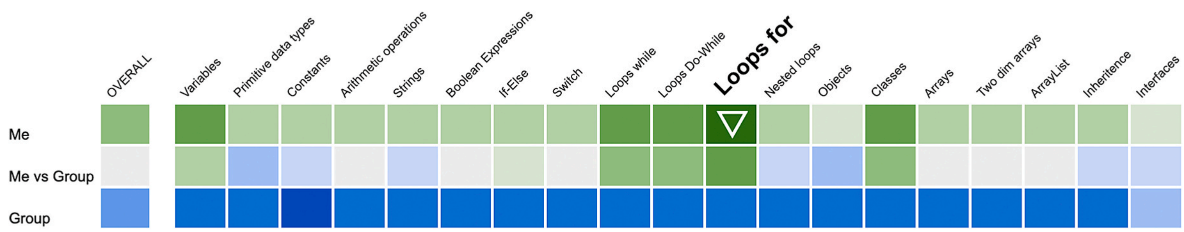

Fig. 2. Example of an Open Learner Model for a Java programming subject. The learner can scrutinise this OLM by clicking any cell. In this figure, the learner clicks *Loops For* to scrutinise the details in Fig. 3.

Fig. 3. Scrutinising the Open Learner Model. This presents details of the learning data that was used to determine the value and corresponding colour gradient in the top level OLM cell. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

mastery-based approach, the teacher may define the level students should achieve by the end of the semester. In that case, the group bar could be replaced with a measure of the standard required at semester end.

There are many other possibilities, such as the student's own intended standard. This has the merit of giving the student more agency. This approach has been supported in systems such as Narcissus (Upton & Kay, 2009). In the discussion section, we will return to this aspect of OLMs and consider alternative approaches.

2.4.2. *Benefits of OLMs*

The AIED literature has now established a substantial body of work that has demonstrated learning benefits from OLMs – for a recent review, see Bull (2020). We now consider the benefits of an OLM, like that in Fig. 2, in the particular ways that it could be valuable in a university subject.

One important benefit is that OLMs support communication between the teacher and students, similar to an *advance organiser* (Ausubel, Novak, Hanesian and others, 1968; Stone, 1983). This is because the OLM presents a brief description of the key aspects the teacher wants their students to learn. The descriptions match the terms used in the T&L activities and assessments. So, at the very start of and indeed at any point in the semester, a student will see these terms as a map of the path ahead.

A second key benefit of the OLM is that it provides the student with a way to *monitor* their learning progress. So, for example, after a series of examples and quizzes, a student could consult their OLM to see whether relevant cells have become green, or greener. If not, they may decide to review the material and redo a quiz. Similarly, a long term OLM could be part of a student's regular *reflection* on their learning progress and then to *plan* longer term strategies. This form of OLM enables the learner to see their growing competence. This can build their confidence that they are progressing in line with their goals. Such opportunities to realise one's competence has been shown to enhance motivation (Patall, Sylvester, & Han, 2014), which is one of the pillars of learning success. This also aligns with the literature on self-efficacy (Bandura, Freeman, & Lightsey, 1999), which highlights the importance of enabling a learner to evaluate their progress towards their goals.

An OLM like Fig. 2 has a third benefit, in making it easier for learners to see how a complex of learning activities and assessments link to the subject's learning goals. So, for example, a skill like "Loops While" in Fig. 2 may be needed for several, perhaps 7 – 12, activities, starting in, say Weeks 2 – 4 and later as other loops are introduced and yet again when "Arrays" are introduced. Even if the student is aware of the aspect of "loops" in each of these, it could be quite difficult to see the big picture of their developing skill on this aspect. Students often struggle to step out of the day-to-day activities to see the big picture. OLMs may help students make this perspective shift.

This type of OLM could also serve as a communication object in several learning contexts. One is where a student meets their tutor for help with a current assignment. With the OLM, the student and tutor together may identify aspects the student needs to develop for the assignment. This could facilitate a discussion about planning how to overcome this before tackling the assignment. Similarly, a student could share their OLM with peers and discuss their progress and how to help each other (Bull & Mabbott, 2006).

There are many potential benefits of OLMs. For example, Bull (2020) lists 16 types of benefits including improved learning, slower knowledge loss, improved self-assessment, increased confidence and greater engagement with the teaching software. There is a growing body of empirical evidence for a complex picture of the benefits of OLMs. They are frequently designed and evaluated for their potential benefits in supporting self-regulated learning and meta-cognitive processes. An early and notable example assessed the OLM in an SQL tutor (Mitrovic & Martin, 2002). The evaluation was in a university subject, where students could volunteer for a between-subjects evaluation study. This found that lower-achieving students in the OLM-group had higher learning gains between the pre-to the post-test than the control group. For higher achieving students, there was no such learning difference but the higher-achiever OLM-group abandoned fewer problems and changed their planning processes more than the control group. A study of meta-cognitive aspects Long and Aleven (2013) compared two OLM conditions in a maths tutoring system: (1) support for self-assessment before seeing the OLM and (2) choices in selecting the level of the next problem. Their 2 × 2 between-subjects study measured learning outcomes (with procedural and conceptual pre- and posts-test tasks) as well as self-assessment accuracy. The OLM conditions improved both, with additional benefits from the version where students self-assessed before seeing their OLM update but problem selection did not give a similar benefit. We present one more example to illustrate the diversity of evaluations. This compared the OLM in Fig. 2 with a version that presented only the individual OLM in the top row. The results of a between-subjects study of in-class use in a Master's Database Management subject showed statistically significant and large benefits in engagement measures, efficiency and learning gain. These examples were chosen for their diverse measures of OLM benefits beyond just the more widely assessed usability and student preferences.

Such uses of an OLM would represent an important shift from the current norm where students may just track their grades on each assessed task. Once an assessment has been completed, this means the student has to accept the grade as immutable. In principle, a low assessment grade, with the grading feedback, should direct the student towards ways to improve their later work in the semester. With suitable design of the assessments, for example, allowing a student to redo a quiz for no credit, the data from such reattempts could be reflected in the OLM. But to truly understand the benefits that might be provided by OLMs we need to dig into the history of learner models (LMs) per se.

2.4.3. A brief introduction to learner models

A core claim of this paper is teachers can use a conceptualisation of learner models to help them design learning data for their subjects. So we need to explain what a learner model is and how a teacher could design one. Although we draw on previous work from AIED, we emphasise that this paper focuses on the design of a learner model that is intended to provide an OLM for students to use in the ways we have described above.

The meaning of the term, learner model, varies across the literature in AIED and Human-computer interaction. Notably, Kay and Kummerfeld (2019) maps out four quite distinct meanings of learner model: the model of the learner in the mind of the teacher; an implicit model in the code of a teaching system; a model of a set of learners; and a model of an individual learner. This last meaning is the one that we now explain.

One recent definition of learner models (Zapata-Rivera & Arslan, 2021) is as: "representations of the learner's knowledge, skills and other attributes … [that] may include cognitive, meta-cognitive, affective, personality, social and perceptual aspects of the learner". Like much AIED work, this definition is for learner models that drive personalisation of a teaching system. AIED has explored many ways to design and build individual learner models. Some are cognitively based, one notable class of these being the ACT-R "cognitive models" based on John Anderson's influential model of learning (Anderson & Lebiere, 2014). Another important class of cognitive learner models is Stellan Ohlsson constraint-based models, based on the theoretical foundations that describes declarative knowledge as constraints and interprets a student's solution to a task in terms of these constraints (Ohlsson, 1994, 2016). These cognitive models have been used in several personalised teaching systems (Mitrovic, 2003) that have been widely deployed and evaluations have shown substantial learning benefits. However, their design is driven by the needs of a teaching system.

Another widely used, but simple, AIED approach to learner modelling is the *overlay model* (Carr & Goldstein, 1977). This is based on a very simple, but very effective, notion: identifying the aspects of learning that are important to model, such as the topics, skills, attitudes etc the teacher wants their students to learn. It may also be useful to model misconceptions – this is useful where the teacher is aware of important and common misconceptions that students bring to their subject.

2.4.4. How are learner models built?

We now briefly describe the key steps AIED system builders take to design learner models, which involves defining:

1. *the learner model ontology*: the *components* of the learner model and the semantic relationships between them;
2. component representation: the way that learning data will be stored and interpreted;
3. data collection: how learning data will be accrued to provide evidence for reasoning about each component.

The learner model ontology is typically a hierarchical collection of components. In Fig. 2, there is just a single list of topics modelled. Another teacher might have combined all three loop components as a single component. Other work, such as Kay (1995); Cook et al. (2015) had a hierarchical knowledge structure, with the OLM initially showing top level components, allowing learners to expand this for more detail and Kay and Kummerfeld (2013) enabled learners to select a component and then display those semantically close to it.

In much of that work, the relationships between components are refinements, meaning that the top level components are related to a set of more detailed refinements of that component. There are also other valuable relationships between components, such as the Piagetian progression of genetic graphs (Goldstein, 1979) which includes other relationships, such as the prerequisite which can be particularly valuable to model. Such rich ontologies have been and continue to be fertile ground for research. But, at present, there are no widely available authoring tools that can make it easy for any teacher to make use of such sophistication.

The design of the representation for each component involves defining the type and value. The most common type is *knowledge*, as in all the components in Fig. 2. Other types that may be useful include the learner's preferences, attitudes, emotions, goals and attributes such as age. This, too, is an aspect where AIED research makes use of many of these to modelling aspects such meta-cognition and affect (Arroyo et al., 2014) or confusion, frustration and "gaming" a system (Arroyo et al., 2014; Richey et al., 2021).

Fig. 2 represents value as a small integer that determines the shade of green to display. However, many other interfaces have been created for OLMs. For example, in an intelligent tutoring system based on sophisticated constraint-based personalised teaching (Bull & Kay, 2016; Mitrovic & Martin, 2002), each component visible in the OLM is a triplet, with measures for:

- correct understanding demonstrated;
- incorrect understanding as indicated by tasks completed;
- the proportion of material not yet covered by the student.

There are many forms an OLM may take. We have limited this treatment to a single, simple example that provides background for our work. The recent and comprehensive review OLMs (Bull, 2020; Bull & Kay, 2016) describes many others. However, despite all of the benefits of having a well-grounded learner model represented as an OLM, one more feature must be considered: if the model cannot be scrutinised and challenged then it will do little to support SRL.

2.4.5. Scrutable OLMs and their learning benefits

An OLM is scrutable[2] if its interface, and the underlying learner model, have been designed so that the learner can use it to answer key questions about the way that the OLM calculates the learner's knowledge level. Essentially, the OLM is an interface element in the student's learning environment. Many AIED systems have the very simple OLMs needed for students to monitor their progress. So, the student can, at a glance, answer questions like: "Am I making progress at the level I want to?". But the underlying learner model could enable students to answer additional questions that as part of reflection, planning and gaining trust in the OLM. For this, the OLM needs to be designed so that the student can scrutinise,[2] the OLM.

Broadly, scrutable OLMs should be designed to follow the visualisation principle of first providing an overview of complex information spaces and then, details on demand (Schneiderman, 1996). In terms of the constructive alignment elements, scrutable OLMs should enable a learner to answer questions such as:

- Why does this system rate my knowledge of X at this level?
- Which T&L activities or assessments determine that rating?
- Can I trust this rating?
- What can I do to improve this rating?

We now explain this in terms of the example in Figs. 2 and 3. In Fig. 2, the learner has clicked the green cell representing *For Loops*, indicated by the triangle in that cell and the large font. This presents the Fig. 3 detail on learning activities used to determine the value of the *For Loops* cell. This means that the learner can judge whether they trust the OLM summary of their learning progress. For example, if they had a lot of help with the quizzes, they might realise that their personal learner model is over-estimating their actual knowledge. Similarly, if there is very little evidence for a component, the student may realise it is not reliable. They could click a quiz cell to go to the quiz and do it so that would provide evidence. Another key aspect of scrutability is the algorithm used to determine the score and so, the shade of green for each component. Higher support for scrutability would enable the learner to discover this.

2.4.6. Summary

We have presented the foundational elements of our approach. Specifically, we have summarised constructive alignment as a way that many tertiary educators currently use, particularly in engineering education. We have also introduced the core concepts that underpin our approach, *learner models* and *OLMs*, with support both for learners to see a depiction of their learning progress and to *scrutinise* the reasoning underlying the OLM.

3. OLMs as drivers for design of learning data, algorithms and AI processes

This section introduces our learner-centred approach to systematic design of learning data by a teacher. We first provide a conceptual overview. Then we work through case studies to illustrate it. We model this description on Biggs' (Biggs, 1999) with a broad conceptualisation followed by illustrative case studies of how to operationalise it. We present our approach as any teacher could apply today, by taking the established AIED concepts of *OLMs*, *learner models* and *scrutability* to provide a systematic way to decide what purpose they want the OLM to serve and then to use that to drive design of data collection and use.

Fig. 4 is the conceptual overview of our approach that starts with the design of an Open Learner Model and then uses this to drive both the learning design and the design of the learning data. The grey boxes are identical to those in Fig. 1 described above. But this figure also shows two classes of learning data, one for the T & L activities and one for the assessments. Our approach calls for the design of this data in parallel with the design (or updating) the actual activities and assessments. We refer to this process as the design of learning data. We now explain all the new elements.

We begin with the design of the OLM. The teacher designs this as the view of learning progress that they want to make available to learners. The teacher needs to design this along with the learning objectives of Biggs' approach. This means that the teacher systematically considers the student view of their learning progress at the same time as they create (or review and update) the usual learning objectives. In addition, the teacher needs to be aware that they will need to define the learner model that will store the data displayed in the OLM.

We now consider the case of the design of the T&L activities and the data to be captured from those. We extend Biggs' constructive alignment, where teachers ensure that these match the intended learning outcomes – now the teacher also designs the learning data that they want to collect from each activity in order to build the learner model so that it can be then shared with the student via the OLM. Now the teacher also specifies the learning data that they require from each activity in order to construct the learner model.

A similar approach applies to the assessments. Of course, it is already common-place to collect assessment data and make that available to students, in the form of marks. It is also common to provide additional feedback to students based on a grading rubric. This may well report grades in terms of a taxonomy such as SOLO or Bloom as described above. The rubric elements may also map to learning objectives.

But a full set of assessments needs to cover multiple learning objectives. Some assessments may be tightly targeted e.g., a short formative quiz on a particular topic. Larger assignments typically involve multiple learning objectives; in this common case, grade data show the student their progress in terms of each learning objective. This is where our approach to the design of learning data aims to systematically align that data to the learning objectives. So, for example, rather than simply track that there is a mark for a particular assessment, the teacher now designs the task and its associated data to align with the learner model they want the students to see. This may mean that, rather than a single mark for this homework, the teacher designs the assessment and grading scheme in such a way that fine grained data works to provide evidence to each relevant part of the learner model.

We now consider the lower bottom box, showing *meta-cognitive activities*. The need for it becomes evident when we consider the nature of the available learning data in relation to learning objectives the teacher considers important. Often the teacher knows that it is either impossible or impractical to collect learning data from either the T&L activities or from assessments. In that case, our approach invites the learner to become a partner in the learner modelling process by answering questions about themselves in a particular class of learning activity created by the teacher. Examples of the types of meta-cognitive activities that could be undertaken by the student include:

- self-rating *knowledge* of a learning objective (also called Feeling Of Knowing), potentially also self-rating their confidence in the rating (Kleitman & Stankov, 2007);
- rating of their *interest* in a learning objective or other aspect related to it, such as the case studies available to study a topic;
- indicating their *attitudes*, where these are among the learning objectives and it is very difficult to model these from with T&L or assessment activities;
- stating their performance *goal*, for example, in terms of a grade or most valued objectives;
- self-rating the importance they of various T&L activities for judging their learning progress – for example about lecture attendance, taking notes while attending lectures, doing self-test activities.

[2] "to examine something very carefully in order to discover information" https://dictionary.cambridge.org/dictionary/english/scrutinize.

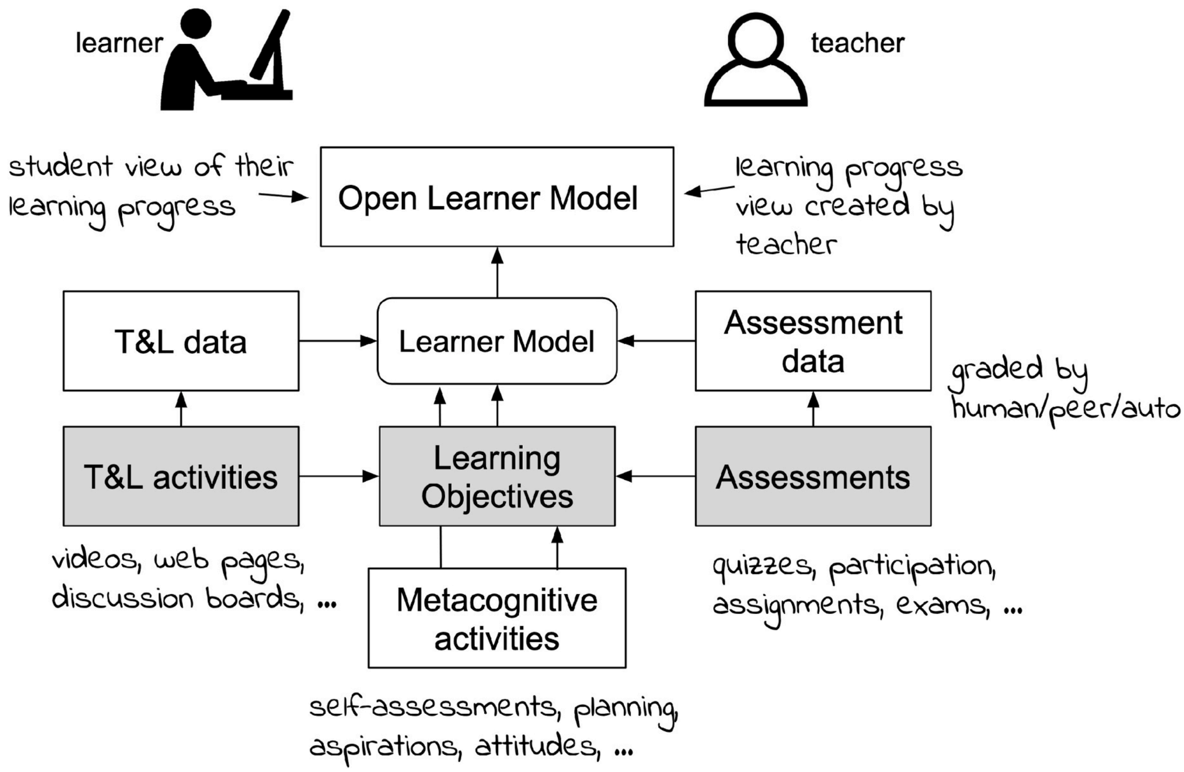

Fig. 4. Overview of OLM-driven approach to both learning design and an aligned design of the learning data.

We separate these from the other grey boxes in the Biggs' constructive alignment for three reasons. First, these activities could be seen as fitting either as T&L activities or as a form of assessment (self-assessment). The second and more important reason is that a focus of our whole approach is for the teacher design the OLM to include what they consider is important. Drawing on the student in this way enables a teacher to still include an important aspect even if other learning data is limited, incomplete, or likely to be noisy and uncertain. The third, even more important reason, is that our learner-centred approach aims to be deeply grounded in giving the learner agency and responsibility.

How can a teacher translate the above conceptualisation into the design of an OLM and its learning data? There are no existing technical tools to do this. This was the case in the mid-1990s when Biggs' constructive alignment papers appeared Biggs (1996). As in Biggs paper, we map out the process and illustrate it with case studies. The high level steps in OLM-driven learning design and aligned design of the learning data are:

1. WHY: Determine the purpose of the OLM (that is, the benefits for students' learning);
2. WHAT: Design the OLM the teacher would like students to use;
3. HOW: Design the learner model ontology;
4. HOW: Design the learning data to provide evidence for the learner model;
5. HOW: Design how the OLM will be used in the learning process';
6. Reflect and refine.

As in most complex design processes, this is an iterative and looping process. It should start with the purposes and move through each of the stages. But in any later stage the teacher may realise that need to backtrack. The next section presents two case studies by two authors of this paper.

4. Case studies to illustrate OLM-driven design of learning data and instructional design

We now illustrate our low-tech approach with two case studies. Fig. 5 gives an overview of the subjects. Key differences are that INFO1112 is a first-year subject, larger, builds on just one technical subject and most importantly, it is very technical compared with DATA3406. Both are very similar in having a complex mix of T&L activities and large and small assessments, very like many science and engineering subjects.

4.1. First case study - large, technical first year subject

We present this case study to elaborate on the first five stages along with specifics of the way we used the approach on INFO1112. For the final reflection stage, we draw on both case at the end of this section.

Step 1 – WHY: Determine the purpose of the OLM

This teacher identified three purposes expressed as questions students should be able to answer from the OLM:

1. How am I doing on mastering each main learning objective?
2. Am I keeping up on *engaging* in all the activities and assessments?
3. Are there activities and assessments that I missed but can still catch up on?

The first question is about students going beyond just the marks for each assessment. This teacher wants to enable students to track their learning in terms of the learning objects. The teacher wanted the second is because their experience made them confident student's learning would benefit from these forms of engagement. The third relates to the subject design, with non-assessed activities such as videos and self-test, no-credit quizzes. The lecturer wanted to encourage students to use these more effectively.

Subject	INFO1112 - Operating Systems and Network Platforms	DATA3406 - Human in the loop Data Analytics
Class level	First year	Third year
Class size	430	160
Assumed knowledge	One programming subject	Two second year subjects (with two first year subjects) - all technical programming and statistics
Nature of this subject	Technical	Multi-disciplinary: ethics, human factors, software engineering team methods, information visualisation, human-computer interaction
T&L activities	Online videos, each with self-test quizzes; live lectures, labs	Lectures, each with formative assessed quizzes and surveys. Tutorials
Assessments	Three large assignments, weekly homework, lab activities and quizzes. No exam.	Two group assignments, weekly lab preparation and participation

Fig. 5. Overview of the two subjects in the case studies.

Step 2 – WHAT: Design the OLM the teacher would like students to use

The teacher considered many ways an OLM could achieve the above purposes. After some exploration of the potential design space, the teacher used a spreadsheet to mock up the design in Fig. 6. This is for a hypothetical student in Week 8 of the semester. At the left are short descriptions of the key OLM components. The first is “Engagement” (in line with the second purpose above). The next is the assumed knowledge from the earlier subject “INFO1110”. The teacher particularly wanted to highlight this so that students would appreciate the importance of the foundations assumed, and it needed for the first purpose. The rest are shorthand descriptions of high level learning objectives that match the formal, published curriculum.

For each week up to Week 8, the OLM shows the student’s progress at one of four levels as described in the legend. An “ideal student” would have scores of 4 for every cell – meaning that at Week 8, the student’s data indicated they were on track for all the OLM components. This student has an “Engagement” level of 3 for each week up to and including Week 8. For the activities in Week 8 alone, their ”Python programming skills” are at Level 2. They could improve to level 3 or 4 if they now complete missed activities (Purpose 3). The teacher’s is intended for students to see fortnightly changes. If this student now completed all available tasks, the OLM would be update at Week 10. White cells mean there were no relevant activities for that learning objective in that week.

Fig. 7 shows the teacher’s design for a student scrutinising the “Engagement” component at the end of Week 8. At the left, this lists the T&L activities and assessment types. The first four, (“Assignment” .. “Homework”) are linked to certain weeks of the semester as shown in coloured cells in the OLM. Grey cells indicate an activity due in the future. The rest occur each week of semester.

	Week 2	Week 4	Week 6	Week 8	Week 10	Week 12	Week 13
Engagement	3	3	3	3			
Assumed from INFO1110	1	1	2	2			
Python programming skills	1	1	2	2			
Unix skills	2	3	3	3			
Data Representation		4	4	4			
Computer Architecture		3	3	3			
Networks and Security				3			
Cloud and Containers							

Legend for level achieved	Explanation
1 - lowest	You have made some progress
2	
3	
4 - highest	You have done everything really well
Not Applicable	
future	

Fig. 6. Top level design of the OLM for a hypothetical student in Week 8. The descriptions at the left are the top level of the learner model ontology. The columns are for fortnightly progress. This student has a mix of levels of progress. The student can scrutinise the detail as illustrated in Fig. 7 and 8.

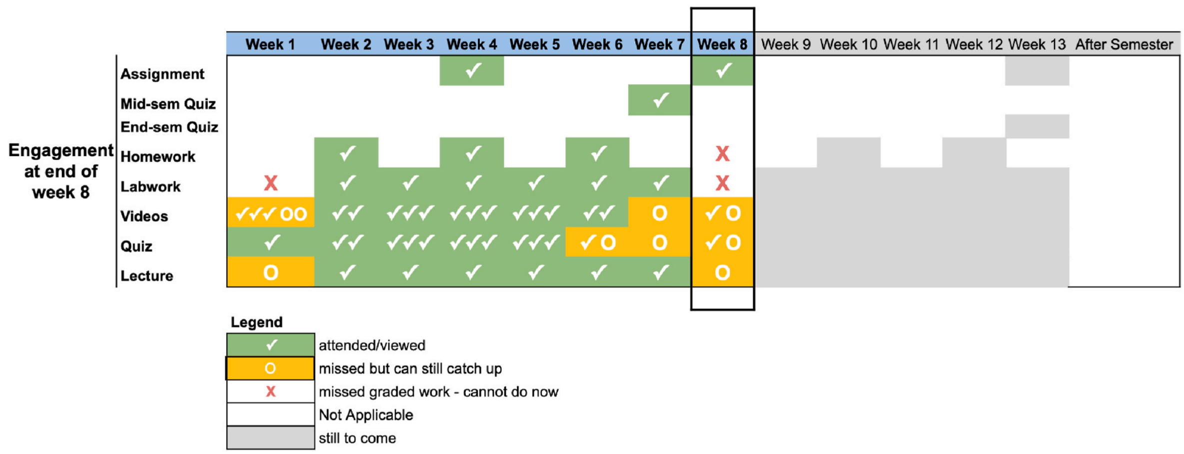

Fig. 7. Example of OLM detail for "Engagement", so that a student can scrutinise the evidence used to determine the value for that component in the top level OLM in Fig. 6.

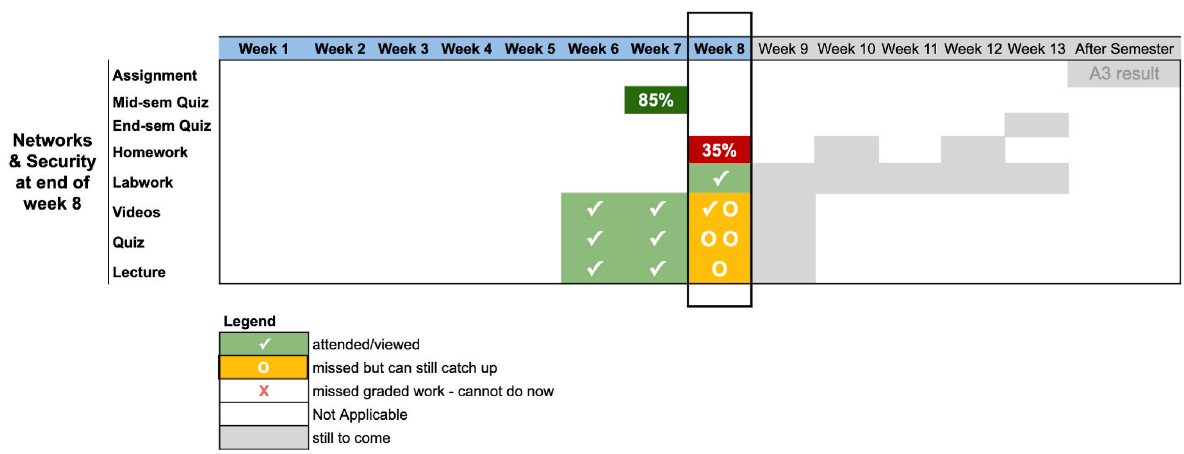

Fig. 8. Example OLM detail, similar to Fig. 7, but now for a *knowledge* component, to support scrutiny of the "Networks and Security" component in Fig. 6.

The last three rows enable a student to see their progress on the ungraded aspects the student can catch up any week (Purpose 3). The student could attend the live lecture or view the recording later. In the OLM, attendance (either live or viewing the video) is shown as a white tick on a green background. When the student does neither, they get "O" on a yellow background. If they then watch the recording by, say, Week 9, this can become a white tick on green in Week 10. Some cells have multiple ticks or "O"s to indicate multiple videos. Other activities, such as the "Labwork", cannot be done after the session has passed. That is indicated by a red X.

Fig. 8 shows a second example, now for the knowledge component, "Networks and Security". The layout is the same as in the last example. A key difference is that the top four activities are graded. "Labwork" is for attendance. Like Fig. 7, the remaining rows are activities where the student can catch up.

As a brief forward reference to Step 5 (Learning Design), this form of scrutable OLM can be created with minimal technology. The teacher can share the spreadsheets and each student can fill it in, by checking through each task. Various class activities could support students in reflecting on their progress, self-assessing it and planning. This could include peer discussions or the tutor discussing these with each student.

Step 3 – HOW: Design the low-tech learner model ontology

In this stage, the teacher designed the *detailed* ontology using a spreadsheet as a low-tech ontology. The spreadsheet makes it easy to document a two-level hierarchy. Fig. 9 shows how each column lists the name of a Learner Model Component, with bold font for the top level components. In the figure, the spreadsheet is expanded to show the four sub-components of "Unix Skills" (Column G). The details of the other top level components are currently hidden. For example, "Data representation" has eight sub-components (in Columns Q .. X, hidden). There are 68 in all. The spreadsheet makes it easy to click on the + to expand out the next level in the ontology or the – to hide that level of detail. This makes it is easy to decide on the level of detail to focus on.

Step 4 – HOW: Design the learning data to provide evidence for the learner model

The teacher then mapped out the potential learning data for each component in the learner model. The rows of the Fig. 9 spreadsheet have the 91 possible evidence sources that were identified from all the T&L activities and the assessments. This part of the spreadsheet also has the top level descriptions of these in bold: three assignments, mid-semester and end-of-semester quizzes, fortnightly homework assignments. This screenshot has hidden the rows with the details of the 13 labs, 28 videos and quizzes and the live lectures. Learning data for each student is obtained from several different systems: Zoom for live lecture attendance, Echo360 for pre-recorded video views, SRES for lab work, Edstem for assignment and homework marks.

Having identified all potential sources of learning data available, the teacher needs to decide just which to use. There will be many factors

	A	B	C	G	K	L	M	N	O	P	Y	AP	BL
1													
2		**Engagement: Completed / Attended / Viewed**	**Assumed from INFO1110**	**Python programming skills**	**Unix skills**	Basic Unix shell commands	shell for loop	shell while loop	shell environment variables	**Data Representation**	**Computer Architecture**	**Networks and Security**	**Cloud and Containers**
3	**Asst 1 (Ed)**	✓	✓	✓	✓	✓							
4	**Asst 2 (Ed)**		✓	✓	✓	✓							
5	**Asst 3 (Ed)**					✓							
6	**Mid Sem Quiz** (Gradescope)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		
7	**End Sem Quiz** (Gradescope)												
8	**Hwk Week 2 (Ed)**	✓	✓	✓	✓	✓	✓			✓			
9	**Hwk Week 4 (Ed)**	✓	✓	✓	✓						✓		
10	**Hwk Week 6 (Ed)**	✓	✓	✓	✓						✓		
11	**Hwk Week 8 (Ed)**												
12	**Hwk Week 10 (Ed)**												
13	**Hwk Week 12 (Ed)**												
14	**Lab work (13 Wks)**	7	-	2	7	✓	✓		✓				
28	**Video segments** (28) (Echo360)	19	-	-	5	✓	✓		✓	4	5		
57	**Self-test quizzes** (26) (Canvas)	19	-	-	5	✓	✓		✓	4	5		
84	**Lecture 13 weeks** (Zoom/video)	7	-	4	4	✓	✓		✓	2	2		
98													
99	91 activities												

Fig. 9. Spreadsheet the teacher gave to students, filled in for an ideal student at the end of Week 7. Learner Model Components are columns. Rows are sources of learning data from assessments and viewing/attending sessions that can contribute to the OLM. The **Unix Skills** in Column K has been expanded. Other bold columns, such as **Data Representation** can also be expanded to show finer detail.

that determine the choices at this step. Some data is very hard to collect or to analyse within pragmatic constraints of real teaching environments. It would be good practice to document this process and the decisions (Wise, Sarmiento and Boothe Jr, 2021).

Fig. 9 shows the OLM spreadsheet filled in for an ideal student at the end of Week 7. The entries in column B indicate the activities completed, attended or viewed. Cells with a tick or number indicate a source of evidence from the row activities for a column component. For example, in row 28 column P the 4 indicates the student viewed four videos with material on Data Representation. Column K, Unix skills, has been expanded to show its sub-components. The ticks in row 28, columns L, M and O indicate the Unix skills covered in those videos.

Step 5 – HOW: Learning Design - classroom use of the OLM

The teacher introduced the OLM to their students during a Week 5 lecture. The lecture started with an overview, introducing the core ideas of the OLM, as explained using the materials above, with the purposes and the design of the OLM. The teacher then explained why the OLM could provide an valuable tool to support students in seeing how the many parts of the subject fit together. After this introduction, an empty version of the Fig. 9 spreadsheet was made available for the students to download their own copy. The teacher then guided the class through the elements of spreadsheet, giving them time to fill it in up to that point of the semester. In later lectures, the lecturer took class time to revisit the spreadsheet.

4.2. *Second case study - 160-student, multi-disciplinary, optional, third-year subject*

We used the earlier case study to illustrate our broad approach. For this case study, we just present the summary of the final design. This teacher defined the purposes of the OLM:

1. Encourage and enable students to track their learning progress on the learning objectives (not just marks on assessments);
2. Help students appreciate the importance of engagement, especially in the COVID enforced online mode, and enable them to self-monitor various aspects of this; and
3. Help students realise the links between topics and why they matter.

The first two are similar to the earlier case study. The third relates to a particular challenge in this subject – most of the topics are linked to each other and in previous years, students had difficulty understanding this.

Fig. 10 shows the spreadsheet form that was introduced in the Week 1 lecture. The first column has the learning components - the high level is in rows 163–169. The detailed learner model ontology is in the first 160 rows. Fig. 11 shows two of these components expanded.

In both figures, Column C is for the knowledge level. The teacher explained this in the first lecture (as part of the pragmatics topic in rows 2 – 11). This was linked to SOLO and Bloom taxonomies and the 3-level classification used in the subject. Column D has a 3-level rating for importance - key for the first purpose. Both these are explained in the README tab. The spreadsheet had data validation in these columns to select the levels. Column B values in Figs. 10 and 11 are defined in the *Comp-type-ev* tab, visible at the bottom of Fig. 10.

Fig. 12 shows the sheet on that tab. Types of evidence are in Column A, with descriptions in Column B. The columns list sources of evidence and green cells indicate where evidence is available for that type of component. The teacher used this in Step 4 to consider what evidence was already available. At the same time, it scaffolded the teacher in considering potential new ways to collect useful evidence. In particular, pink Column D is for evidence that could be collected in the in-lecture surveys each week. This process drove the design of survey questions about aspects where no other data was available from the previous version of the subject.

The OLM was integrated into many elements of the teaching. The first lecture introduced the spreadsheet and explained how the OLM would be used throughout the semester, in many of the weekly lab preparation mini-assignments and in the exam. Each lecture started with a slide listing the detailed components to be introduced - the slides were available for students to download. The lecturer discussed this OLM ontology, noting links to the current and up-coming practical work. In some weeks, the lecturer screen-shared the spreadsheet to highlight links between the topics previously covered and to the high-level topics that pervade the subject - *0. Learning in this subject* and *1. People - high level*. In-class surveys at the start of the lectures often asked for students to self-report their knowledge, experiences and attitudes. The design of many of these was driven by the OLM – so that these could provide students and the teacher with evidence about the OLM components relevant for starting that class and complementing data from T&L activities and assessments.

About half the weekly assignments asked students to write a multiple-choice question that covered one or two specified OLM topics or linked them to practical work. In the corresponding tutorial class, students worked in small groups to share these and then each group shared a co-created one with the class, followed by class discussion. The

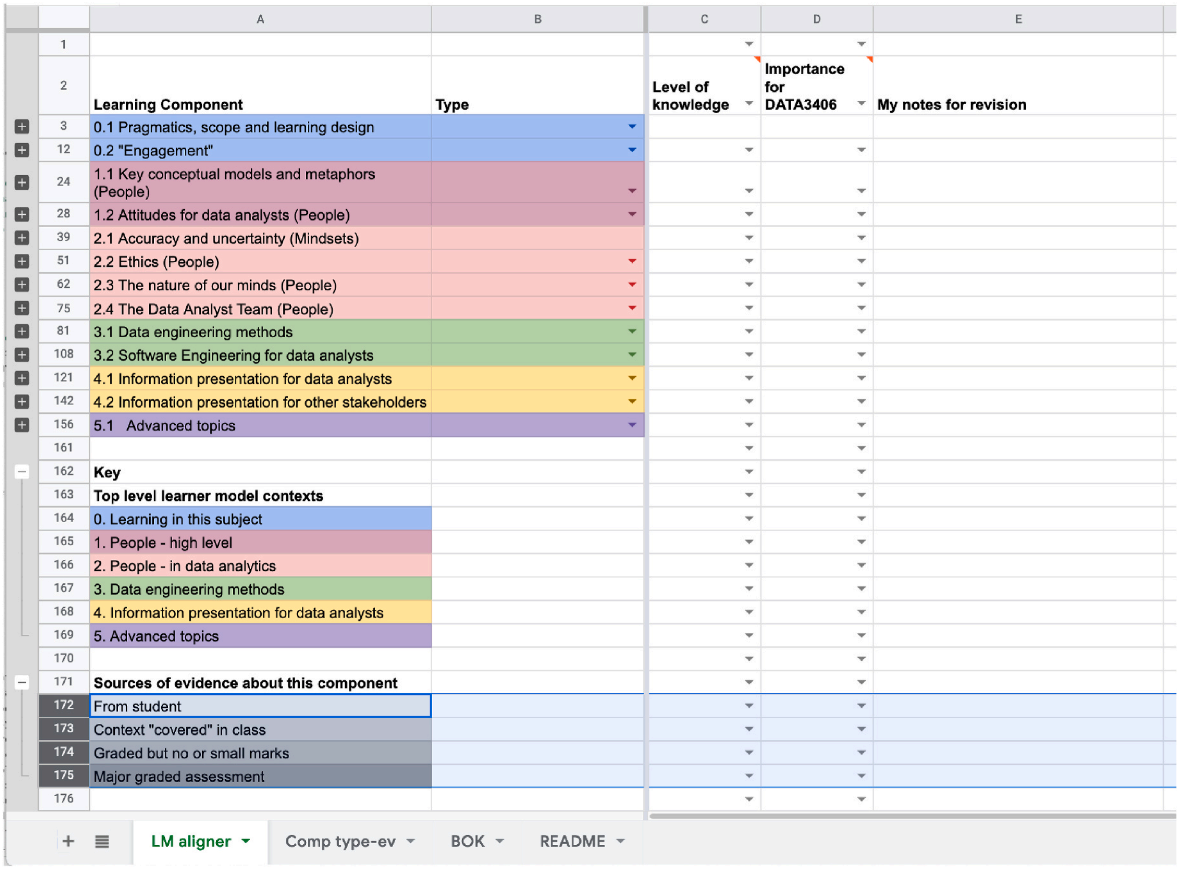

Fig. 10. Top level ontology for the DATA Learner Model. The key, expanded, shows the colour coding of the top level topics. The sources of evidence are described in the grey scales. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

exam had similar multiple choice questions. Students were encouraged to use the OLM spreadsheet in their revision, to map topics across the subjects T&L activities and assessments. They could, optionally upload their sheet with their exam. Fifteen percent of the class did this and about half of these submissions had enough content to be a useful reference for students. About a quarter were quite rich, showing links across the subject and comments about student perspectives. These provided valuable insights for the lecturer. In discussions with the consultative class group (students from each class who volunteered to meet the lecturer most weeks to share their views and experiences), the use of the OLM was refined and student comments indicated how it helped them see the cross links in the subject.

Key features of this case study are that the OLM perspective transformed the teaching in several ways. It drove the learning design to use the OLM as a means to improve communication of the purpose of each lecture and the links across all aspects of the subject. It started communication about aspects like engagement. It led the lecturer to create many of the questions for in-lecture surveys and quizzes to provide new evidence. The detailed components served as a form of glossary or definition of the actual meanings of the learning objectives. It provided a vocabulary for students to refer to variable aspects of the subject such as in lectures, discussion boards and the consultative group. Future refinements will include ways for students to co-create the OLM. For example, tutorial preparation could include defining detailed links to students' other subjects or current news and these could be shared in the tutorials. This would build on comments from the consultative group about links to their other subjects such as law, philosophy and psychology, all with valuable and often complementary perspectives.

Step 6 – Reflect and refine

Both case studies enabled the teachers create new ways to communicate with students. Our low-tech approach enabled the teachers to use the OLM to drive the design of learning data and learning activities and consider how these intersect. The spreadsheets for incorporating the OLM into the learning design. The approach made for extreme flexibility in the choice of data sources. This was a form of Wizard-of-Oz approach to creating the learner model, with the students as the wizards updating the OLM spreadsheets. Reflecting on the timescales for computer supported constructive alignment, our case studies match the approaches Biggs used in the mid-1990s; even now, tools like CUSP, U-Planner, Akari and CourseLoop are becoming more widely deployed. We now turn to ways technology can support the six steps of our approach.

5. How to build an OLM using an existing learning analytics platform

One of the objectives of this paper is to demonstrate that this is achievable for instructors using existing learning technologies. Fig. 13 demonstrates how the OLM envisaged in Fig. 8 can be implemented using the Student Relationship Engagement System (SRES), a learning analytics platform that allows a significant level of instructor-driven customisation (Arthars & Liu, 2020). Through SRES, a teacher could design a live OLM that directly represents progress (Fig. 13-A) and also instructor-curated suggestions that are personalised (Fig. 13-C) according to student progress and goals. Metacognitive activities can be integrated into the OLM, where learners set goals, reflect, and plan in response to simple prompts (Fig. 13 B and D). To allow this, SRES gives instructors the ability to curate, manipulate, and analyse data from a range of sources. Data can be synchronised from the LMS, or can be entered directly into SRES by teaching assistants and students themselves through instructor-defined forms, or can be ingested through an

	A	B
1		
2	**Learning Component**	**Type**
3	0.1 Pragmatics, scope and learning design	
12	0.2 "Engagement"	
13	Lecturer academic interests	Engagement (emotional)
14	Tutor academic interests	Engagement (emotional)
15	Shared own academic interests	Engagement (emotional)
16	Consistently do homework on time	Engagement (behavioural)
17	Consistently attend lectures	Engagement (behavioural)
18	Consistently attend tutes	Engagement (behavioural)
19	Consistently monitor ed	Engagement (behavioural)
20	Team engagement - with asst group	Engagement (emotional)
21	Class engagement - with tute group and whole class in class activities, breakouts	Engagement (emotional)
22	Answer other student questions on Ed	Engagement (cognitive)
23	Post comments in lectures (Ed/Zoom)	Engagement (cognitive)
24	1.1 Key conceptual models and metaphors (People)	
25	Conceptual model: **Humans** and the **loops** of data analytics	Knowledge (new)
26	Conceptual model: human cognitive, affective and social aspects -- two elephant **metaphors**	Knowledge (new)
27	Conceptual model: **disciplines** of Ethics, Software Engineering, HCI, visualisation	Knowledge (new)
28	1.2 Attitudes for data analysts (People)	
39	2.1 Accuracy and uncertainty (Mindsets)	
51	2.2 Ethics (People)	

Fig. 11. Expansion the learner model ontology to show details of two high level components. This shows the different forms of engagement. Rows 24–27 has the expanded details of a knowledge component.

API. Once data are curated in one place within SRES, instructors can design adaptive web pages and messages that will be automatically personalised to the needs of individual learners based on the data available in its database.

An instructor would first determine what is made available to learners through the OLM (Fig. 14, label A at the left). For example, activity-level representations of progress are shown (Fig. 14 A1), allowing learners to see progress towards goals and what activities are yet to be completed, as well as personalised suggestions (Fig. 14 A3). In keeping with our parallel focus on SRL, meta-cognitive prompts to learners are also displayed and responses stored (Fig. 14 A2 and A4). Teaching, learning, and assessment activities (B) are then designed, many of which can provide evidence for the learner model: knowledge evidence derives from regular assessments, homework, lab work, and tests; engagement evidence derives from participation in live lectures, lab work (reported by teaching assistants using the SRES web app, Fig. 14 B2), homework, self-test quizzes, and self-monitoring of confidence in understanding material in the recorded videos (reported by learners through the SRES LTI, Fig. 14 B1).

These data points, or evidence for the learner model, are curated by an instructor into SRES. Each subject has a separate instructor-customised database within SRES (Fig. 14 C), where columns hold data about individual learners (in rows), similar to a spreadsheet. In a similar way, instructors can define internal calculations to manipulate data within SRES and calculate the value of each component of the OLM. For example, video viewing data and Feeling of Knowing self-reports can be aggregated into a visual representation. At the next level of abstraction an engagement score, for example, can be calculated from the submission status of assessments, video views and self-monitoring, as well as quiz and homework completions - these are marked with ‘1’ in Fig. 14.

This real-world application of an OLM provides learners and instructors with a communication object for discussions of progress, provides learners with a form of advance organiser and personalised feedback to plan and monitor progress, and can help promote metacognition through SRL prompts. For the instructor, all data, visualisations, and the OLM design itself are built through a single platform, SRES.

6. Discussion and research agenda

This section discusses key ways that our OLM-driven data design can be used now as well as the research needed to make it easier for teachers to use. We have organised this section in terms of three key stakeholders in this process. First, we focus on student perspectives, then for teachers, we consider issues ranging from learning design that makes use of OLMs to the pragmatics. We then move from what the broader issues for researchers in education, AIED and learning analytics.

6.1. Student perspectives

OLMs as a communication object. Pivotal to the concept of an OLM is that they can facilitate the communication between teacher(s) and learner and also can be compared and discussed between learners (Bull & Mabbott, 2006). From the perspective of communication, the OLM can play two roles: it can be the communication channel and it can be

	A	B	C	D	E	F
1			Sources of evidence			
2	Types	Description	Graded Task	Lecture survey	In-lecture quiz	Action logs eg Ed, Zoom survey....
3	Knowledge (new)	What students learn in this subject				
4	Knowledge (assumed)	What should know before this subject				
5	Misconception	Misconceptions often seen in this subject				
6	Attitude	Attitudes this subject aims to develop				
7	Preference	Student preferences relevant to this subject				
8	Attribute	Student attributes				
9	Engagement (behavioural)	"Attends" learning activities , e.g., accesses to Canvas pages, Ed discussion board				
10	Engagement (cognitive)	Make effort to understand ideas and master skills				
11	Engagement (emotional)	Interest, enthusiasm, absence of anger, anxiety, and boredom				
12	Engagement (agentic)	Strives to personalise learning for own goals				
13	Goal	Student aspirations				
14						
15	Legend					
16	Data directly available					
17	Student self-assessed					
18						

Fig. 12. Evidence types: The names in Column 1 determine the values in the LM-Aligner tab. Columns C – F show the evidence available. This is part of step 4 on determining the evidence available.

the object of the communication. In this paper, we treated the OLM mainly as the channel for teacher-learner communication, as a tool for the teacher to provide orientation and express expectations and for the learner to communicate information over and above what is displayed in the standard learning environment via regular teaching tools such as the LMS. Seen from the perspective of Activity Theory (Engeström, 2014). This means that two activity systems – one made up by the students and their goals, the other by the teacher and their goals – use the same tool, the OLM. Tensions, if not contradictions, can arise when the goals are different and also because of differences in capacity to use the OLM: it may be a routine tool to the teacher but not to the student. There is a small body of research on OLMs as communication objects, such as reviewed in (Bull, 2020). One question is how communication around a student's OLM is similar to and different from discussions around other forms of Learning Analytics dashboards and different from discussions around other learning materials, such as e-portfolios.

OLMs to help students understand and manage their learning. One key message in our work is that the terms OLM, learner model, ontology and scrutability, can be a valuable part of the vocabulary of teachers but they could also be valuable for students. This is because they can serve as way to think about learning and learning progress, based on evidence from learning data. This can be part of shifting students' conversations and thinking from marks to a richer view of the whole learning picture in terms of their developing knowledge, skills and attitudes. In this paper we have focused on OLMs that concentrate on the provision of information regarding topic mastery. However, it is easy to imagine using the OLM-design framework proffered here to provide information regarding key components of self-regulated learning such as effort regulation (known to be strongly related to achievement, (Richardson, Abraham, & Bond, 2012), perhaps in conjunction with tools to assist in assignment planning (e.g. (Kia, Teasley, Hatala, Karabenick, & Kay, 2020)).

OLMS as dashboard for fast-use and slow and purposeful use. From a student perspective, OLMs, with support for scrutiny of the underlying learner model, are a particular class of dashboard that has roles for both fast-use monitoring of the high level OLM and slow-use during scrutiny (Kay, Rus, Zapata-Rivera, & Durlach, 2020). Much work on dashboards has emphasised the notion that they provide information "at a glance" (Schwendimann et al., 2016). The OLM literature has many examples of on-screen skill meters that fit this use for fast self-monitoring. Our top level OLM in Fig. 6 could serve this role if it were automatically updated after each action by a learner. The OLMs we can easily build now, as described in our case study, are for slow and deliberate thinking. In a flexible system, like SRES showcased here, an OLM can encourage meta-cognitive reflection, and be accompanied by personalised messages regarding possible next steps for students, in line with their goals or interests. Framed in this way, OLMs provide "actionable information", an emphasis within the Learning Analytics community.

Assess the actual benefits of this approach for learners. The core goal of our work is to improve learners' experiences and outcomes. It is challenging to assess this. It also involves both learners and teachers. One approach is for teachers to draw on data about subject-level student perceptions and evaluations of learning, as has been reported in evaluation studies for the benefits of constructive alignment (for example, (Larkin & Richardson, 2013)). Of course, these are limited; the two subjects in our case studies had large increases in student ratings and final grades of the subjects but it is impossible to attribute these to the OLM use since the teachers also changed many other aspects from the previous year, in line with normal refinement of teaching based on teacher reflection. Early work on *instructional alignment* at the task level showed large learning benefits (Cohen, 1987) - our work operates at the semester level. So, we need to establish systematic ways to collect meaningful measures of learners' engagement, their perceptions and learning outcomes, based on carefully designed assessments of aspects that the OLM was intended to support. These should also draw on qualitative evidence from teaching teams. All these need to be linked to the ways that the teacher used our approach.

6.2. Teacher perspectives - thinking about curriculum and data design

Maintaining a learner-centred and learning-centred focus on the data design process. The starting point for our OLM-driven data design is the teacher's definition of the *goals* for the OLM. Part of the teacher's

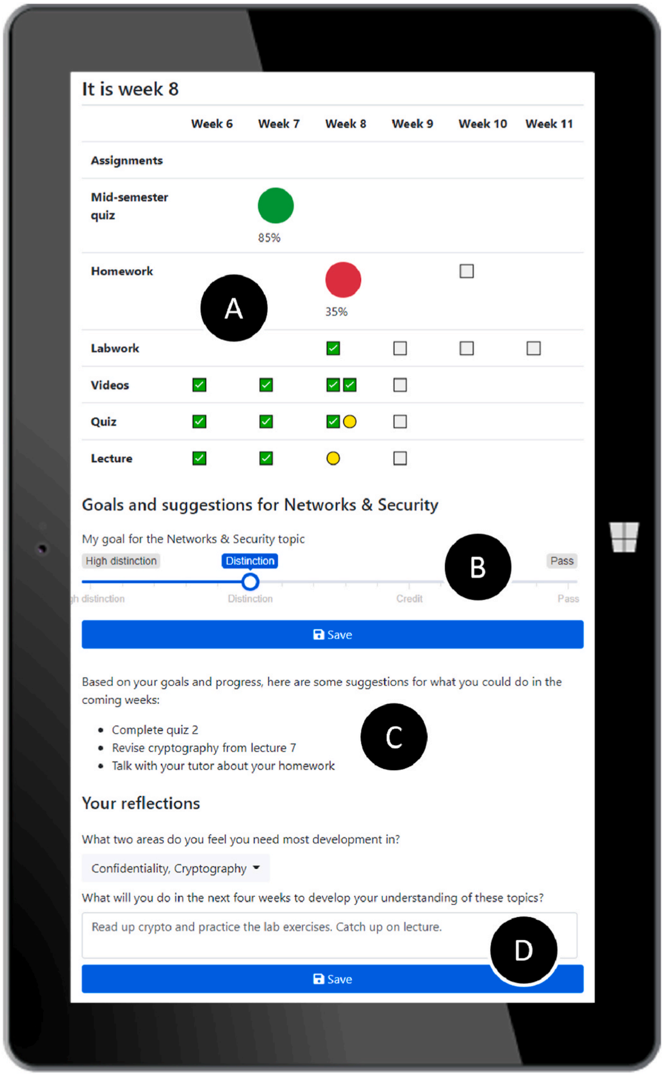

Fig. 13. An OLM implemented within an existing learning analytics platform. Student progress is presented (A) alongside instructor-customised feedback (C) that is based on learner-set goals (B) and connected with reflective prompts (D).

reflective practice calls for assessing whether those goals have been met. A very simple and valuable way to do this is in a simple survey. It could start with a general question, such as “How useful did you find the OLM?”. Additional questions should ask more specific ones about the driving goals for the design. For example, in our case study, the lecturer wanted to communicate the importance of engagement; a survey could ask questions about these aspects as well as open questions for qualitative evidence. In a platform like SRES, the teacher could integrate this into the OLM interface. It is also easy and useful to instrument the OLM to gather data about its use.

Making use of the OLMs in teaching. One key reason to create OLMs and integrate them into the teaching has been described in (Bjork, Dunlosky, & Kornell, 2013): “people often have a faulty mental model of how they learn and remember, making them prone to both misassessing and mismanaging their own learning.” To really tackle this, it is critical that teachers incorporate the use of the OLMs into T&L activities. This can be part of building students’ assessment literacy (O’Donovan, 2017; Price, Rust, O’Donovan, Handley, & Bryant, 2012). The learning design should help students to understand the meaning and purpose of the OLM in their subject. It also calls for class time for use of the OLMs, helping students learn how and when to scrutinise and how to make best use of the details.

OLMs that go beyond knowledge components. OLMs that link tightly with formal learning objectives may mainly show knowledge components. This is in line with much AIED literature on OLMs. Our case study also had engagement, albeit only behavioural engagement, similarly to much learning analytics work such as (Papanikolaou, 2014). Other aspects may also be valuable, including of cognitive and affective engagement (Fredricks, Blumenfeld, & Paris, 2004) as well as agentic engagement (D’Mello, 2021). If the teacher values student engagement, the OLM provides an opportunity to communicate that to the students. Our OLM-driven approach can then be a driver for designing the T & L activities shown in Fig. 4. There are also other important learning goals that could be modelled, such as attitudes and skills like group collaboration.

Care in the use of peer comparators. Our introduction to OLMs showed a student’s progress compared to a peer group. In most AIED OLMs, this is not done; rather the individual’s own progress is displayed. This owes much to the powerful influence of the Cognitive Tutors which were deeply committed to mastery learning (Corbett & Anderson, 1994). But it is important to discuss the risk of OLMs that do present a comparison of a student’s performance with that of a peer group. Care is needed to avoid potential negative effects of visualisations on students, and the differential impact they may have on subgroups of students depending on their motivation beliefs. Within a class there will be a range of motivation profiles, for example the mix of perceived competence, goal orientations, and task values found by (Linnenbrink-Garcia et al., 2018). Goal orientation theory typically defines three orientations: mastery, performance-approach, and performance-avoidance; and it is included as a component of SRL. Higher levels of achievement are associated with students with high levels of mastery-approach, or high levels of both mastery and performance-approach goals; whereas, there is a negative association with performance-avoidance goals (Niemivirta, Pulkka, Tapola, & Tuominen, 2019). Performance-avoidance goals are also related to maladaptive outcomes such as help-avoidance and self-handicapping (Senko & Tropiano, 2016). There is also emerging evidence that although students with high levels of both mastery and performance-approach goals achieve well, well-being indicators may be negative (e.g. (Tuominen, Juntunen, & Niemivirta, 2020)). This makes the question of whether it is beneficial to present comparative information via an OLM or Learning Analytics dashboard to students a very real issue, especially given the relative technical ease with which it can be provided. A recent study of students’ use of a collection of three LA dashboards found that high SRL students viewed the comparison dashboard the least, and low SRL students the most (Kia et al., 2020). This is an unhelpful outcome for this subgroup of students. A challenge for the future research agenda may be to think how we can design visualisations that promote a mastery orientation.

Differences between institutional constructive alignment and OLM-driven data and learning design. OLM-driven data design needs to fit within the institutional context. Section 4 showed the differences between institutional or degree-level constructive alignment and our approach. Clearly these need to be consistent but such constructive alignment needs to have modest numbers of high level learning objectives. Tools like CUSP, U-Planner and Akari have a key role in supporting degree-level approval of these. They are then published and frozen for the semester. By contrast, OLM-driven data design is fine-grained (i.e. more low level) and can allow the teacher flexibility to make changes during the semester. Much learning technology is also supported at an institutional level. In the future, institutions may require that those tools make it far easier for teachers to make flexible use of learning data as needed for OLM-driven data design.

Pragmatics of how to do OLM-driven data design today. The core goal of this paper is to create a way for teachers to think about the design of learning data. As in our case study, the design process may include some consideration of rich OLMs. There is so much that a teacher may like students to be aware of. But pragmatics may mean that it is currently only practical to create quite modest, low-tech learner models and

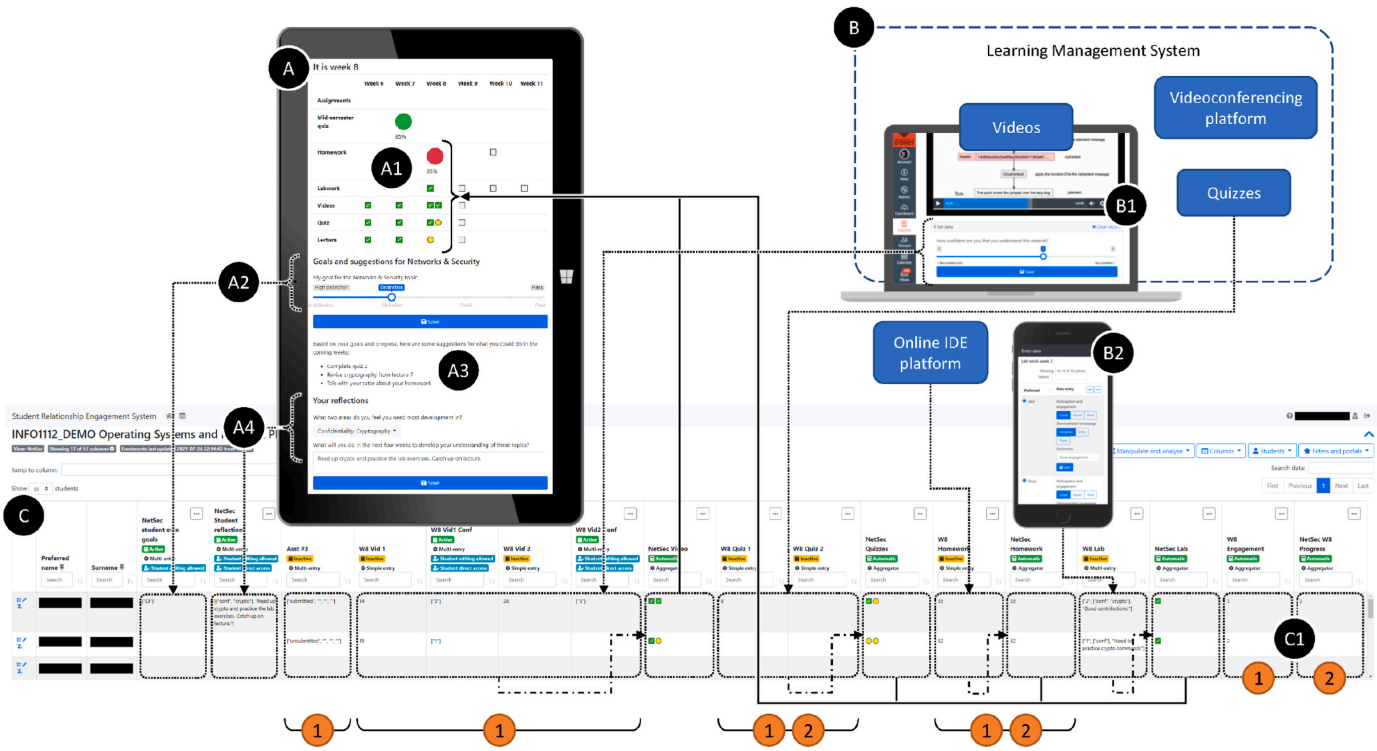

Fig. 14. A learner model and SRL-enriched OLM built using the Student Relationship Engagement System (SRES). Learners access an OLM (A) that represents their learning progress, which derives data from teaching, learning, and assessment activities (B) that are curated in the SRES database (C) by instructors.

OLMs. Our case studies emphasised aspects that can be done now. But we now turn to our research agenda that should make it easier to do more and to gain insights about how to do this effectively.

6.3. Researcher perspectives and research agenda

In this final section, we take a step back, working to sketch out a position that would help the research agenda involving the design and use of scrutable OLMs to move forward. This is a rich field, with decades of research behind it, but it covers a number of different avenues, which can make it difficult to draw all of the relevant threads together. We hope that this discussion will help those who would like to contribute to this exciting field as the tools that we need to support it start to achieve maturity.

Fig. 15 is an overview of the 'how' steps, both for the low-tech approach described in the case studies and a future with richer support from technology. The figure shows the potential for richer OLM ontologies and glossary tools, such as Apted, Kay, and Lum (2004); Zablith and Azad (2021). Our spreadsheet approach supported a two-level hierarchy. A more complex spreadsheet could easily support a deeper hierarchy but the use would be more complex. Our case studies showed how the spreadsheet such as in Fig. 12 supports systematic consideration of the potential evidence sources already available and new ones to be created. For automated data harvesting, an interface should enable a teacher to specify the place that learning data can be accessed so that it can then be automatically linked to a tool like SRES. The third part of the figure is for interfaces for the individual student to use. These may be similar to the many current learner model interfaces Bull (2020). This is a rich area for future research and development, particularly with scaffolding for learners to systematically work through meta-cognitive processes such as reflection and planning. We now discuss the state of the art and barriers for creating these tools.

The first point to note is that, even with the advances in technology represented by new systems such as SRES, it is still very difficult for a teacher to develop an OLM that is automatically updated, based on the

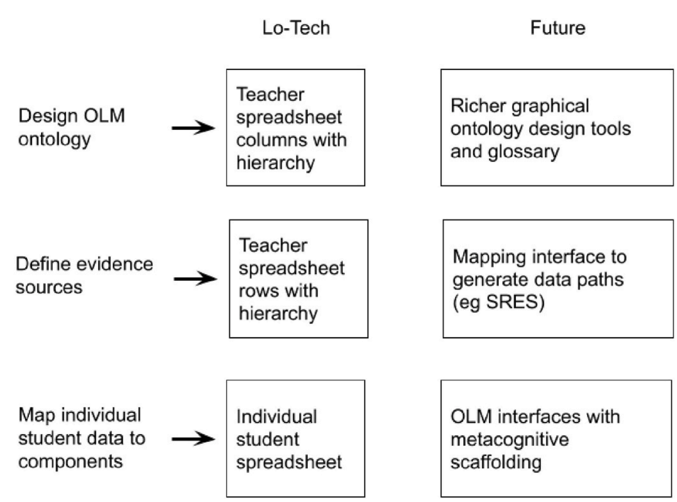

Fig. 15. Left: Overview of the three stages of our approach. Middle: maps this to the low-tech approach used in our case studies. Right: The future directions with improved tools for the teacher and the learner.

evidence available about each component in the OLM. One of the biggest problems concerns the complexity of the learning system in which they participate. As we have seen in Section 4, even individual subjects in large tertiary institutions make use of a wide array of tools in delivering a rich learning experience, which creates a number of problems for those who would build learner models over diverse forms of data. For example when attempting to build a Learner Model across multiple sources, two events that are highly similar might be represented in a markedly different manner; the LMS might store separate marks for each quiz response, while a multiple choice test inside a video delivered by a LTI integration may have been designed to store only the total mark across all questions asked in the video. Similarly, one source might use JSON as its data format and the other XML. And the data might be sent to very different locations. Even with an advanced ETL process, it often becomes necessary to somehow map between these divergent sources before the data can be used. Some solutions to this problem have emerged over the years. For example, Bull, Johnson, Alotaibi, Byrne, and Cierniak (2013) manually created an OLM from multiple data sources, including use of a simple competency-based OLM, Facebook and chat discussions, demonstrating that this is, in principle, possible. Similarly, Kitto et al. (2017) provides an overview of the *Connected Learning Analytics* (CLA) toolkit, which aggregated data from various social media (e.g. Facebook, Slack, GitHub and Twitter) to produce student facing dashboards that were used to provoke student reflection in their participation in class based learning activities. However, we are yet to see solutions emerge that make it easy for a non-technical instructor to merge all relevant data streams into an OLM – or even to include some of these streams.

While proponents of the semantic web suggested that this problem of *educational data interoperability* should be addressed almost from the beginning of the movement (Aroyo & Dicheva, 2004; Kump, Seifert, Beham, Lindstaedt, & Ley, 2012), no general solution emerged. The most recent iteration of this ongoing program of work has led to the creation of new activity data standards such as xAPI and IMS Caliper, which were proposed as a way to unify data from a wide range of tools and environments. Unfortunately the dream has yet to be realised in a systematic manner, and indeed, the fact that two standards exist for the same problem could be seen as contributing substantially to its ongoing nature (Kitto, Whitmer, Silvers, & Webb, 2020). Ongoing research and development will be required to achieve a seamless integration between the rich variety of tools used to deliver online learning, and there is a risk that every institution will "reinvent the wheel" while no general solution is provided by vendors. Collaborations between institutions and across sectors will help to avoid this duplication.

Better data pipelines, able to integrate data from both endorsed university systems and from other tools that students nominate (e.g. social media and other software that students choose to use to monitor progress, set goals, determine deadlines, etc) would enable the construction of richer learner models with a more complete representation of student's knowledge, skills and attributes. However, data in this format is far more difficult to process. It requires a move beyond a simple counting of events, and towards methods that analyse social interactions (Shum & Ferguson, 2012), perform quantitative ethnography (Shaffer, 2017) and use Natural Language Processing (NLP) (McNamara, Allen, Crossley, Dascalu, & Perret, 2017). The interfaces built over complex data with rich learner models would help to promote student reflection in class based activities, with an associated potential increase in the capability of students to reflect generally. The OLM could work to provide a boundary object, which supports discussions between teaching staff and students. However, such tools require guidance and support to ensure their wider take up (Van Merriënboer & Kirschner, 2012; Zapata-Rivera & Greer, 2002). It is important to realise that it is sometimes possible to provide this guidance in an indirect manner, for example by embedding an OLM in a learning design that supports students in making sense of the model (Knight, Gibson, & Shibani, 2020). More work in required in this area to understand the potential affordances of using OLMs in a class based setting.

Nonetheless, analysing the cognitive constructs behind the clicks of learning analytics (Wise, Knight, & Buckingham Shum,) is currently an extremely challenging problem. The approach proposed here would require that a teacher, who may not be expert in the fields used to generate the results, be able to extract learning data, and then use it to generate learner models and then present them to students in an interpretable manner. This is an extremely challenging problem at present. Research is required to bridge this important gap.

Of course, along with richer learner data comes a wide range of ethical issues around Fairness, Accountability and Trust (FAT). While the concept of privacy has been widely covered by the field of LA (Pargman & McGrath, 2021), FAT has been relatively absent until quite recently, but is starting to gain a substantial profile in the field (Tsai, Perrotta, & Gašević, 2020; Wise, Sarmiento, & Boothe, 2021), along with recent calls to encourage LA practitioners to develop a keener ethical lens when building their tools (Johanes & Thille, 2019; Kitto & Knight, 2019). With the OLM as a driver for designing learning data, the teacher can keep the student's perspective in mind so helping to encourage the field in this direction. The notion of scrutability can help the teacher ensure that they provide the student with information to just the fairness of the data interpretation encapsulated in the OLM.

This now brings us to an important point that is often missed across the literature: over the decades a wide range of different fields completed work in the learning sciences, represented by the AIED, as well as Educational Data Mining (EDM), Computer Supported Collaborative Learning (CSCL), Learning at Scale (L@S), and Learning Analytics (LAK) communities. This means that we frequently see research performed by one sub-community missed by another. For example, the early work on OLMs was missed in the Bodily et al. (2018) review of learner facing dashboards, despite the clear overlap between the two fields. A challenge of terminology presents, where concepts such as learner model, student model and user model are variously adopted across different fields, with an associated siloing of the results from each. A substantive research effort is required to draw the core ideas from these different fields together and to harmonise the different approaches that they use for evaluating the rigour of the different approaches and methodologies springing from each of them.

Tying many of the above themes together, with richer learner models, that make use of interoperable learner data from a wide range of systems there is potential to construct OLMs over rich lifelong learner models, that help people to monitor and plan their learning over a lifetime of change and disruption (Kay & Kummerfeld, 2019). Achieving such a goal would help to support an ongoing partnership between humans and AI assistants, drawing on the particular power of each (Baker, 2016). Even without achieving this grand aim, we have shown how a typical teacher of a university subject could incorporate these ideas into the learning and data design. This is a form of AIED evolution (Roll & Wylie, 2016) that would see the field move towards a more human centred perspective that empowers people in an age that increasingly reliant upon AI.

7. Conclusions

The work on constructive alignment (Biggs, 1996, 1999) has been valuable at several levels. We now focus on two of them that inspired our work. First, it provides the term that teachers can use to talk with each other about their learning design. Beyond that, the availability of a term for an important idea is an aid for the teacher to think about that idea, in line with the Sapir-Whorf Hypothesis (Pinker, 2003; Sapir, 1929; Whorf and others, 1940). One of our goals is to present the AIED terms *OLM*, *learner model* and *scrutability* as tools to help teachers think about and discuss their learning design processes.

Biggs' papers on constructive alignment also made an important contribution because they taught by example how a teacher could readily take this idea and use it to design their own subjects. Similarly, one of our goals has been to show that teachers can, today, take these ideas and follow a learning design process that enables them to create OLMs and to use the goal of creating them as a driver for the many decisions involved in designing the learning data for their own subject.

One very important difference between the context of Biggs and current education is the huge growth in use of technology for learning. It had such a small role at that time. Today, it is pervasive, with a typical university subject making use of multiple tools, notably LMSs and either within or outside them, tools for automated grading, group discussions and chat, self-tests and many specialised software tools relevant to a particular subject. This makes it timely to extend the simple but powerful ideas of constructive alignment, adding intellectual tools for teachers to use in designing the learning data for their subjects.

A second, even more important difference is that Biggs' work was teacher centred, where ours is deeply student centred. It is based on making the OLM a boundary object that the teacher creates as a form of communication with students. To make it serve that role, the teacher crafts the description of the learning goals, aiming to make this meaningful for students in the context of their subject. And where constructive alignment should help the teacher be systematic in designing assessments, our approach can be seen to operationalise that in the design of the learning data for the OLM.

In this paper, we have outlined ways to take these core notions of OLM, learner model and scrutability from AIED and apply them to learning design that incorporates data design. We have also outlined the ways teachers can do this today. We have also outlines the potential ways that could provide richer, automated ways to create OLMs that draw on diverse forms of learning data that come from multiple learning tools and offer scaffolding to assist each learner in building skills in self-regulated learning that is based on both self-perceptions and other learning data as evidence of learning progress.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Anderson, L. W. (2002). Curricular alignment: A re-examination. *Theory Into Practice, 41*, 255–260. Publisher: Taylor & Francis.

others Anderson, L. W., & Bloom, B. S. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.

Anderson, J. R., & Lebiere, C. J. (2014). *The atomic components of thought*. Psychology Press.

Apted, T., Kay, J., & Lum, A. (2004). Supporting metadata creation with an ontology built from an extensible dictionary. In *International conference on adaptive hypermedia and adaptive web-based systems* (pp. 4–13). Springer.

Aroyo, L., & Dicheva, D. (2004). The new challenges for e-learning: The educational semantic web. *Journal of Educational Technology & Society, 7*, 59–69.

Arroyo, I., Woolf, B. P., Burelson, W., Muldner, K., Rai, D., & Tai, M. (2014). A multimedia adaptive tutoring system for mathematics that addresses cognition, metacognition and affect. *International Journal of Artificial Intelligence in Education, 24*, 387–426. Publisher: Springer.

Arthars, N., & Liu, D. Y. T. (2020). How and why faculty adopt learning analytics. In *Adoption of data analytics in higher education learning and teaching* (pp. 201–220). Springer.

others Ausubel, D. P., Novak, J. D., & Hanesian, H. (1968). *Educational psychology: A cognitive view*. New York: Publisher: Holt, Rinehart and Winston.

Azevedo, R. (2020). Reflections on the field of metacognition: Issues, challenges, and opportunities. *Metacognition and Learning, 15*, 91–98. Publisher: Springer.

Baker, R. S. (2016). Stupid tutoring systems, intelligent humans. *International Journal of Artificial Intelligence in Education, 26*, 600–614. Publisher: Springer.

Bakharia, A., Corrin, L., De Barba, P., Kennedy, G., Gašević, D., Mulder, R., et al. (2016). A conceptual framework linking learning design with learning analytics. In *Proceedings of the sixth international conference on learning analytics & knowledge* (pp. 329–338).

Bandura, A. (2005). The evolution of social cognitive theory. In K. G. Smith, & M. A. Hitt (Eds.), *Great minds in management* (pp. 9–35). Oxford University Press.

Bandura, A., Freeman, W., & Lightsey, R. (1999). *Self-efficacy: The exercise of control*.

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education, 32*, 347–364.

Biggs, J. (1999). What the student does: Teaching for enhanced learning. *Higher Education Research and Development, 18*, 57–75.

Biggs, J. B. (2011). *Teaching for quality learning at university: What the student does*. UK): McGraw-hill education.

Biggs, J. B., & Collis, K. F. (2014). *Evaluating the quality of learning: The SOLO taxonomy (Structure of the observed learning outcome)*. Academic Press.

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology, 64*, 417–444. Publisher: Annual Reviews.

Bloom, B. S. (1974). Time and learning. *American Psychologist, 29*, 682. Publisher: American Psychological Association.

Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher, 13*, 4–16. Publisher: Sage Publications Sage CA: Thousand Oaks, CA.

Bodily, R., Kay, J., Aleven, V., Davis, D., Jivet, I., Xhakaj, F., et al. (2018). Open learner models and learning analytics dashboards: A systematic review. In *Proceedings of the eight international learning analytics & knowledge conference* (pp. 1–10). Tex. organization: ACM.

Brusilovsky, P., Somyürek, S., Guerra, J., Hosseini, R., & Zadorozhny, V. (2015). The value of social: Comparing open student modeling and open social student modeling. In *International conference on user modeling, adaptation, and personalization* (pp. 44–55). Springer.

Bull, S. (2020). There are open learner models about. *IEEE Transactions on Learning Technologies, 13*, 425–448.

Bull, S., Brna, P., & Pain, H. (1995). EXtending the scope of the student model. *User Modeling and User-Adapted Interaction, 5*, 45–65. Publisher: Springer.

Bull, S., Johnson, M. D., Alotaibi, M., Byrne, W., & Cierniak, G. (2013). Visualising multiple data sources in an independent open learner model. In *International conference on artificial intelligence in education* (pp. 199–208). Springer.

Bull, S., & Kay, J. (2013). Open learner models as drivers for metacognitive processes. In *International handbook of metacognition and learning technologies* (pp. 349–365). Springer.

Bull, S., & Kay, J. (2016). SMILI: A framework for interfaces to learning data in open learner models, learning analytics and related fields. URL: *International Journal Artificial Intelligence in Education, 26*, 293–331. https://doi.org/10.1007/s40593-015-0090-8. tex.bibsource: dblp computer science bibliography http://dblp.org tex.bib. url http://dblp.org/rec/bib/journals/aiedu/BullK16 tex.timestamp. Sun, 28 May 2017 13:19:04 +0200.

Bull, S., & Mabbott, A. (2006). 20000 inspections of a domain-independent open learner model with individual and comparison views. URL: *Intelligent Tutoring Systems, 4053*, 422–432. publisher: Springer http://www.springerlink.com/content/u131742343825504/.

Carr, B., & Goldstein, I. P. (1977). *Overlays: A theory of modelling for computer aided instruction*. Massachusetts Inst of Tech Cambridge Artificial Intelligence Lab. Technical Report.

Cohen, S. A. (1987). Instructional alignment: Searching for a magic bullet. *Educational Researcher, 16*, 16–20. Publisher: Sage Publications Sage CA: Thousand Oaks, CA.

Cook, R., & Kay, J. (1994). The justified user model: A viewable, explained user model. In A. Kobsa, & D. Litman (Eds.), *Fourth international conference on user modeling UM94 (proceedings)* (pp. 145–150). MITRE, UM Inc.

Cook, R., Kay, J., & Kummerfeld, B. (2015). MOOClm: User modelling for MOOCs. In *International conference on user modeling, adaptation, and personalization* (pp. 80–91). Springer.

Corbett, A., & Anderson, J. (1994). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction, 4*, 253–278. Publisher: Springer.

Corno, L. (2008). On teaching adaptively. *Educational Psychologist, 43*, 161–173. Publisher: Taylor & Francis.

Deeley, S. J., & Bovill, C. (2015). Staff student partnership in assessment: Enhancing assessment literacy through democratic practices. *Assessment & Evaluation in Higher Education, 42*, 463–477. Publisher: Taylor & Francis.

D'Mello, S. K. (2021). *Improving student engagement in and with digital learning technologies. OECD digital education outlook 2021 pushing the frontiers with artificial intelligence, blockchain and robots: Pushing the frontiers with artificial intelligence, blockchain and robots* (Vol. 79). OECD Publishing.

Engeström, Y. (2014). Activity theory and learning at work. In *Tätigkeit-Aneignung-Bildung* (pp. 67–96). Springer.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*, 59–109. Publisher: Sage Publications Sage CA: Thousand Oaks, CA.

Goldstein, I. P. (1979). The genetic graph: A representation for the evolution of procedural knowledge. *International Journal of Man-Machine Studies, 11*, 51–77. Publisher: Elsevier.

Hanus, M. D., & Fox, J. (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers & Education, 80*, 152–161. Publisher: Elsevier.

Johanes, P., & Thille, C. (2019). The heart of educational data infrastructures= Conscious humanity and scientific responsibility, not infinite data and limitless experimentation. *British Journal of Educational Technology, 50*, 2959–2973. Publisher: Wiley Online Library.

Kay, J. (1995). The um toolkit for cooperative user modelling. *User Modeling and User-Adapted Interaction, 4*, 149–196. Publisher: Springer.

Kay, J., & Kummerfeld, B. (2013). Creating personalized systems that people can scrutinize and control: Drivers, principles and experience. *ACM Transactions on Interactive Intelligent Systems (TiiS), 2*, 1–42. Publisher: ACM New York, NY, USA.

Kay, J., & Kummerfeld, B. (2019). From data to personal user models for life-long, life-wide learners. *British Journal of Educational Technology, 50*, 2871–2884. Publisher: Wiley Online Library.

Kay, J., & Lum, A. (2005). In *Exploiting readily available web data for scrutable student models* (pp. 338–345). AIED.

Kay, J., Rus, V., Zapata-Rivera, D., & Durlach, P. (2020). Open learner model visualizations for contexts where learners think fast or slow. In *Design recommendations for intelligent tutoring systems: Data visualization* (pp. 125–136). Publisher: US Army Combat Capabilities Development Command – Soldier Center, ISBN 978-0-9977257-8-0.

Khan, I., & Pardo, A. (2016). Data2U: Scalable real time student feedback in active learning environments. In *Proceedings of the sixth international conference on learning analytics & knowledge* (pp. 249–253).

Kia, F. S., Teasley, S. D., Hatala, M., Karabenick, S. A., & Kay, M. (2020). How patterns of students dashboard use are related to their achievement and self-regulatory engagement. In *Proceedings of the tenth international conference on learning analytics & knowledge* (pp. 340–349).

Kitto, K., & Knight, S. (2019). Practical ethics for building learning analytics. *British Journal of Educational Technology, 50*, 2855–2870. Publisher: Wiley Online Library.

Kitto, K., Lupton, M., Davis, K., & Waters, Z. (2017). Designing for student-facing learning analytics. *Australasian Journal of Educational Technology, 33*.
Kitto, K., Whitmer, J., Silvers, A., & Webb, J. (2020). *Creating data for learning analytics ecosystems*. URL: Society for Learning Analytics Research (SoLAR). SoLAR Position Paper https://www.solaresearch.org/wp-content/uploads/2020/09/SoLAR_Position-Paper_2020_09.pdf.
Kleitman, S., & Stankov, L. (2007). Self-confidence and metacognitive processes. *Learning and Individual Differences, 17*, 161–173. Publisher: Elsevier.
Knight, S., Gibson, A., & Shibani, A. (2020). Implementing learning analytics for learning impact: Taking tools to task. *The Internet and Higher Education, 45*, Article 100729. Publisher: Elsevier.
Kump, B., Seifert, C., Beham, G., Lindstaedt, S. N., & Ley, T. (2012). Seeing what the system thinks you know: Visualizing evidence in an open learner model. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 153–157).
Larkin, H., & Richardson, B. (2013). Creating high challenge/high support academic environments through constructive alignment: Student outcomes. *Teaching in Higher Education, 18*, 192–204. Publisher: Taylor & Francis.
Linnenbrink-Garcia, L., Wormington, S. V., Snyder, K. E., Riggsbee, J., Perez, T., Ben-Eliyahu, A., et al. (2018). Multiple pathways to success: An examination of integrative motivational profiles among upper elementary and college students. *Journal of Educational Psychology, 110*, 1026. Publisher: American Psychological Association.
Long, Y., & Aleven, V. (2013). Supporting students' self-regulated learning with an open learner model in a linear equation tutor. In *International conference on artificial intelligence in education* (pp. 219–228). Springer.
Marton, F., Hounsell, D., & Entwistle, N. J. (1997). *The experience of learning: Implications for teaching and studying in higher education*. Scottish Academic Press.
McNamara, D. S., Allen, L., Crossley, S., Dascalu, M., & Perret, C. A. (2017). *Natural language processing and learning analytics. Handbook of learning analytics 93*. Publisher: Society for Learning Analytics Research.
Mitrovic, A. (2003). An intelligent SQL tutor on the web. *International Journal of Artificial Intelligence in Education, 13*, 173–197. Publisher: IOS Press.
Mitrovic, A., & Martin, B. (2002). Evaluating the effects of open student models on learning. In *International conference on adaptive hypermedia and adaptive web-based systems* (pp. 296–305). Springer.
Niemivirta, M., Pulkka, A. T., Tapola, A., & Tuominen, H. (2019). Achievement goal orientations: A person-oriented approach. In *The cambridge handbook of motivation and learning* (pp. 566–616). Cambridge University Press.
O'Donovan, B. (2017). How student beliefs about knowledge and knowing influence their satisfaction with assessment and feedback. *Higher Education, 74*, 617–633. Publisher: Springer.
Ohlsson, S. (1994). Constraint-based student modeling. In *Student modelling: The key to individualized knowledge-based instruction* (pp. 167–189). Springer.
Ohlsson, S. (2016). Constraint-based modeling: From cognitive theory to computer tutoring–and back again. *International Journal of Artificial Intelligence in Education, 26*, 457–473. Publisher: Springer.
Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology, 8*, 422. Publisher: Frontiers.
Papanikolaou, K. A. (2014). Constructing interpretative views of learners' interaction behavior in an open learner model. *IEEE Transactions on Learning Technologies, 8*, 201–214. Publisher: IEEE.
Pargman, T. C., & McGrath, C. (2021). Mapping the terrain of ethics in learning analytics: A systematic literature review of empirical research. *Journal of Learning Analytics, 8* (2), 123–139.
Park, O., & Lee, J. (2004). Adaptive instructional systems. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (pp. 469–484). Taylor & Francis.
Patall, E. A., Sylvester, B. J., & Han, C.w. (2014). The role of competence in the effects of choice on motivation. *Journal of Experimental Social Psychology, 50*, 27–44. Publisher: Elsevier.
Pinker, S. (2003). *The language instinct: How the mind creates language*. Penguin UK.
Pintrich, P. R. (2000). Chapter 14 - the role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–502). San Diego: Academic Press. https://doi.org/10.1016/B978-012109890-2/50043-3. URL: https://www.sciencedirect.com/science/article/pii/B9780121098902500433.
Popham, W. J. (2008). *Transformative assessment*. Association for Supervision and Curriculum Development.
Price, M., Rust, C., O'Donovan, B., Handley, K., & Bryant, R. (2012). *Assessment literacy: The foundation for improving student learning*. Oxford Centre for Staff and Learning Development, Oxford Brookes University.
Reigeluth, C. M., Myers, R. D., & Lee, D. (2016). The learner-centered paradigm of Education (Chapter 1). In *Instructional-design theories and models* (Vol. IV, pp. 5–32). Routledge. The learner-c entered paradigm of education.
Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*, 353. Publisher: American Psychological Association.
Richey, J. E., Zhang, J., Das, R., Andres-Bray, J. M., Scruggs, R., Mogessie, M., et al. (2021). Gaming and confrustion explain learning advantages for a math digital learning game. In *International conference on artificial intelligence in education* (pp. 342–355). Springer.
Roll, I., & Wylie, R. (2016). Evolution and revolution in artificial intelligence in education. *International Journal of Artificial Intelligence in Education, 26*, 582–599. Publisher: Springer.
Sapir, E. (1929). *The status of linguistics as a science. Language , 207–214*. JSTOR.
Schneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. In *Proc. Of the 1996 IEEE symposium on visual languages* (pp. 336–343). Washington, DC: IEEE Computer Society.
Schwendimann, B. A., Rodriguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., et al. (2016). Perceiving learning at a glance: A systematic literature review of learning dashboard research. *IEEE Transactions on Learning Technologies, 10*, 30–41. Publisher: IEEE.
Self, J. A. (1990). Bypassing the intractable problem of student modelling. *Intelligent tutoring systems: At the crossroads of artificial intelligence and education, 41*, 1–26. Publisher: Norwood, NJ: Ablex.
Self, J. (1999). Open sesame?: Fifteen variations on the theme of openness in learning environments. *International Journal of Artificial Intelligence in Education, 10*, 1020–1029.
Senko, C., & Tropiano, K. L. (2016). Comparing three models of achievement goals: Goal orientations, goal standards, and goal complexes. *Journal of Educational Psychology, 108*, 1178. Publisher: American Psychological Association.
Shaffer, D. W. (2017). *Quantitative ethnography*. Lulu. com.
Shum, S. B., & Ferguson, R. (2012). Social learning analytics. *Journal of educational technology & society, 15*, 3–26. Publisher: JSTOR.
Stone, C. L. (1983). A meta-analysis of advance organizer studies. *The Journal of Experimental Education, 51*, 194–199. Publisher: Taylor & Francis.
Tsai, Y. S., Perrotta, C., & Gašević, D. (2020). Empowering learners with personalised learning approaches? Agency, equity and transparency in the context of learning analytics. *Assessment & Evaluation in Higher Education, 45*, 554–567. Publisher: Taylor & Francis.
Tuominen, H., Juntunen, H., & Niemivirta, M. (2020). Striving for success but at what cost? Subject-specific achievement goal orientation profiles, perceived cost, and academic well-being. *Frontiers in Psychology, 11*, 2470. Publisher: Frontiers.
Upton, K., & Kay, J. (2009). Narcissus: Interactive activity mirror for small groups. In *UMAP09, user modeling, adaptation and personalisation* (pp. 54–65).
Van Merriënboer, J. J., & Kirschner, P. A. (2012). *Ten steps to complex learning: A systematic approach to four-component instructional design*. Routledge.
Wahlster, W., & Kobsa, A. (1989). User models in dialog systems. *User models in dialog systems, 4–34*. Publisher: Springer.
others Whorf, B. L. (1940). *Science and linguistics*. IN, USA: Bobbs-Merrill Indianapolis.
Wise, A., Knight, S., Buckingham Shum, S., . Collaborative learning analytics., in: Cress, U., Rosé, C., Wise, A., Oshima, J. (Eds.), International handbook of computer-supported collaborative learning. volume Vol. 50, p. in press. Publisher: Springer.
Wise, A. F., Sarmiento, J. P., & Boothe, M., Jr. (2021). Subversive learning analytics. In *LAK21: 11th international learning analytics and knowledge conference* (pp. 639–645).
Zablith, F., & Azad, B. (2021). *Reconciling instructors' and students' course overlap perspectives via linked data visualization*. IEEE Transactions on Learning Technologies Publisher: IEEE.
Zapata-Rivera, D., & Arslan, B. (2021). Enhancing personalization by integrating top-down and bottom-up approaches to learner modeling. In *International conference on human-computer interaction* (pp. 234–246). Springer.
Zapata-Rivera, D., & Greer, J. E. (2002). Exploring various guidance mechanisms to support interaction with inspectable learner models. In *International conference on intelligent tutoring systems* (pp. 442–452). Springer.
Zimmerman, B. J., & Moylan, A. R. (2009). Self-regulation: Where metacognition and motivation intersect. In *Handbook of metacognition in education* (pp. 299–315). Routledge/Taylor & Francis Group.