Escape from the Skinner Box: the case for contemporary intelligent learning environments

Benedict du Boulay
Human Centred Technology Group
School of Engineering and Informatics
University of Sussex
Brighton
BN1 9QJ
UK

Email: B.du-Boulay@sussex.ac.uk
Telelphone: 07761614354

Biography
Benedict du Boulay is an Emeritus Professor of Artificial Intelligence in the School of Engineering and Informatics at the University of Sussex and Visiting Professor at University College London. His main research area are the applications of Artificial Intelligence in Education, particularly the issues around modelling and developing students' metacognition and motivation. He was President (2015-2017) and is currently Treasurer and Secretary of the International Society for Artificial Intelligence in Education, and an Associate Editor of its International Journal of Artificial Intelligence in Education.

Abstract
Intelligent Tutoring systems (ITSs) and Intelligent Learning Environments (ILEs) have been developed and evaluated over the last 40 years. Recent meta-analyses show that they perform well enough to act as effective classroom assistants under the guidance of a human teacher. Despite this success, they have been criticised as embodying a retrograde behaviourist technology. They have also been caught up in broader controversies about the role of Artificial Intelligence (AI) in society and about the entry of big data companies into the education market and the harvesting of learner data. This paper concentrates on rebutting the criticisms of the pedagogy of ITSs and ILEs. It offers examples of how a much wider range of pedagogies are available than their critics claim. These wider pedagogies operate at both the screen level of individual systems, as well as at the classroom level within which the systems are orchestrated by the teacher. It argues that there are many ways that such systems can be integrated by the teacher into the overall experience of a class. Taken together, the screen-level and orchestration-level dramatically enlarge the range of pedagogies beyond what was possible with the "Skinner Box".

Practitioner Notes

What is already known about this topic

- A variety of ITSs and ILEs exist and are commercially successful, particularly in the USA.
- ITSs and ILEs have typically been used one-to-one with a single student.
- Early ITSs and ILEs had a reputation for promoting passive learning.

What this paper adds

- ITSs and ILEs have been evaluated to give better post-test learning outcomes than a single teacher in a conventional classroom or lecture theatre, but not as good as a skilled human teacher working one-to-one.
- Support tools for teachers are now available to help them manage a class of students using ITSs and ILEs
- ITSs and ILEs are not limited to isolated use by individual students.
- The pedagogy at screen-level should be distinguished from the pedagogy deriving from the way the system(s) are orchestrated in the classroom.

Implications for practice and/or policy

- The best results are obtained by involving and training teachers in the effective orchestration of the ITSs and ILEs in their classroom.
- Paying attention to effective orchestration has positive benefits in terms of learning behaviours and outcomes.
- ITSs and ILEs can now perform as effective classroom assistants.

Introduction

From their earliest days, Intelligent Tutoring Systems (ITSs) and Intelligent Learning Environments (ILEs) have been concerned to mimic expert human teaching and to develop novel pedagogies that exploit the specific affordances of these environments (see, e.g. Burton & Brown, 1979, page 14; Clancey, 1979; O'Shea, 1979). Since then, the pedagogies have developed from the one-to-one, skill-focused tutoring of some early intelligent tutoring systems through to dialogue-based conceptual interactions that may take the learner's affective state in account, and on to an explicit expectation that these systems will typically be used in a blended fashion, and that the teachers who deploy them need to be supported to orchestrate their use.

ITSs and ILEs use artificial intelligence techniques to adapt and scaffold the experience of the individual learner in ways that attempt to maximise the quality of their learning and/or minimise their learning time. For example, a recent evaluation of the Assistments system found that "the use of Assistments caused 75% more learning than in a typical year" (Roschelle, Feng, Murphy, & Mason, 2016). Such AI-based systems achieve this by building a model of the learner based on their interaction with the system in conjunction with a model of the domain. For the sake of simplicity, the rest of the paper will refer to all such systems under the generic term "Intelligent Learning Environments" (ILEs).

Within the context of formal education, seven meta-reviews have compared the quality of the learning outcomes with ILES as compared to using more traditional methods (Kulik & Fletcher, 2016; Ma, Adesope, Nesbit, & Liu, 2014; Nesbit, Adesope, Liu, & Ma, 2014; Steenbergen-Hu & Cooper, 2013, 2014; VanLehn, 2011). Ma et al. (2014), for example, in their meta-analysis report:

> "The use of ITS was associated with greater achievement in comparison with teacher-led, large-group instruction (g = .42), non-ITS computer-based instruction (g = .57), and textbooks or workbooks (g = .35). There was no significant difference between learning from ITS and learning from individualized human tutoring (g = –.11) or small-group instruction (g = .05). Significant, positive mean effect sizes were found regardless of whether the ITS was used as the principal means of instruction, a supplement to teacher-led instruction, an integral component of teacher-led instruction, or an aid to homework." (Ma et al., 2014, page 901)

Taken together, these seven meta-reviews involved more than 180 comparative studies, though there may be some double counting. Overall they show that such systems performed better with an effect size of 0.47, in the narrow sense of learning outcomes, than a teacher or lecturer working with a whole class, but not as well as a skilled human tutor working one-to-one with an individual learner with an effect size of -0.19 (for a summary of these meta-reviews, see du Boulay, 2016).

Despite these positive results, ILE's have come under a number of criticisms. These fall into three general types. First, there are criticisms of their inherent pedagogy and a believed lack of solid evaluation of effectiveness. Second, there are criticisms of the way ILEs are currently being deployed in some educational environments. Third, there are concerns about the role of AI and big companies in education and the harvesting of learner data. For example, Herold (2017) has criticised personalised learning in general, and ILEs in particular, on a number of grounds. These include, inclarity about the meaning of the term "personalised learning", a weak evidence base for their effectiveness, attempts by some proponents to fill the whole of the student's day with screen-based activity, and the argument that the pedagogy within these systems is inadequate and redolent of earlier, discredited, behaviourist traditions.

Pedagogy and evaluation

In terms of their pedagogy, ILEs have been labelled as the "new behaviourism", offering a contemporary version of Skinner Boxes (Watters, 2015, 2017; Wilson & Scott, 2017). A "Skinner box"(Skinner, 1968) was an electro-mechanical teaching system embodying the following principles:

> "In acquiring complex behavior the student must pass through a carefully designed sequence of steps, often of considerable length. Each step must be so small that it can always be taken, yet in taking it the student moves somewhat closer to fully competent behavior. The machine must make sure that these steps are taken in a carefully prescribed order" (Watters, 2015).

Watters (2015) argues that:

> "[It's] not to say that the influence of Skinner and behaviorism are gone. Far from it. Behaviorism has persisted - although often unnamed and un-theorized - in much of the technology industry, as well as in education technology – in Turing machines not simply in teaching machines."

Wilson and Scott (2017) state that "most ITSs basically follow one pedagogical method in which the learner passively accepts the information given by the teacher or instructional media". They also claim that:

> "ITSs by their nature follow "instructional" paradigms, where the logical constraints of the procedural knowledge to be mastered demand rigid imposition of sequences of didactic demonstration and associated equally constrained assessed tasks." (Wilson & Scott, 2017, page 14)

This is wrong on two counts. First, as I shall show, ILEs can teach both conceptual as well as procedural knowledge. Second, as I shall also show, they do not necessarily demand rigid sequences of didactic demonstration. Wilson and Scott, in turn, cite Bloom (1995), who identifies as one of three key problems with ITS: "The users are not willing to accept the ITS, because of the limited pedagogy that is available, for example, there may not be enough feedback from the "teacher" in the ITS."

This notion of limited pedagogy and lack of feedback had some truth in terms of the very early days of Intelligent Tutoring Systems (ITSs), a sub-category of ILEs, but later systems, as explained later in this paper, offer many different kinds of feedback both at the domain level as well as at the metacognitive and affective levels.

Even earlier than C. P. Bloom (1995), Rosenberg (1987) criticised the ILEs of that time as not well grounded in theories of learning and not having been evaluated via convincing control group studies. Again, allowing for these criticisms to be partially correct in 1987, they no longer hold more than 30 years later, see the earlier remarks about detailed evaluations. The main thrust of the criticisms about pedagogy is that the learner has little agency and is coerced along a particular, though individualised, learning path. While this may have been true of some specific kinds of ILE, it is not a general feature. In the first place, some contemporary ILEs leave the agency about what to do next with the learner, and in the second place, some ILEs deliberately involve the human teacher in the way that the ILE is managed and used in the classroom. Other systems are adaptive not just to the degree of understanding and skill development of the learner, but also to their affective disposition and motivation. Finally, ILEs typically provide (and adapt) feedback at different levels of granularity, as well as offering both domain-level and metacognitive help when learners get stuck or are not going about their learning in an effective manner. The body of this paper provides details of such systems.

Methods of deployment

The second kind of criticism is the way that some organisations have chosen to deploy ILEs. In the worst cases, student sit all day at lap-tops, albeit using adaptive systems, but with insufficient involvement of human teachers, insufficient meaningful interactions with their peers, and insufficient time away from their screens.

This dystopian vision, is a legitimate criticism of the way that some ILEs are deployed, but not of the qualities of the ILEs themselves. It implies a future where learners are trapped for many hours a day at screens, offering an impoverished and dehumanised interaction to the benefit of big tech companies. This vision is not always shared by learners, for instance, when there were demonstrations by students about the Summit Learning Program:

> "Unfortunately we didn’t have a good experience using the program, which requires hours of classroom time sitting in front of computers. Not all students would receive computers, the assignments are boring, and it’s too easy to pass and even cheat on the assessments. Students feel as if they are not learning anything and that the program isn’t preparing them for the Regents exams they need to pass to graduate. Most importantly, the entire program eliminates much of the human interaction, teacher support, and discussion and debate with our peers that we need in order to improve our critical thinking." (Tabor, 2018)

This paper argues for a different vision: one where education is still human-centred, where human teachers support the general direction of learning, including helping learners “learn how to learn”, and are confident that they can offload some of their work to ILEs operating as classroom assistants, making use of varied pedagogies.

The role of AI companies

A different kind of criticism arises from the increasing, and scary, use of AI in many areas of life and has led to major concerns such as (i) the dehumanisation of teaching and the sacking of teachers and (ii) the potential take-over of the educational system by corporate interests in order to harvest the data generated by learners – taking them a further step along the already well developed path from “learner” to “customer” to themselves being the “product” via the harvesting of their data (see the chater by Williamson in this special issue, as well as Williamson, 2018a, 2018b; Williamson, 2018c). Protests by parents, teachers and students in the UK and the USA have highlighted these concerns (see for example, https://twitter.com/AGavrielatos/status/1121704316069236739 and https://twitter.com/hashtag/TellPearson?src=hash).

The issues of the corporate takeover of education and data harvesting are not addressed here, but are amply covered by Petr & Thille and Kitto & Knight in their chapters in this special issue.

This paper's defence of the richness of ILE pedagogy is conducted by illustrating a number of pedagogies incorporated into ILEs directly, or indirectly via the way they are deployed, that go far beyond behaviourist principles. These include (i) adaptively assisting the learner to work through examples rather than solving problems, (ii) adaptively assisting the learner to learn by teaching, (iii) conducting a dialogue with the learner about the concepts to be learned, (iv) cultivating metacognitive reflection about their learning strategies, (v) adaptively helping learners to work in pairs.

A number of mainstream examples of ILEs are described briefly to illustrate the main and varied thrusts of their pedagogy. In general, these systems use AI to create a model of the learner in order to choose what best next task to present and/or how to frame feedback and advice at various levels of granularity. Some also use AI to conduct a dialogue in either written or spoken English, and possibly to infer something of the learner's affective state. Many involve the use of one or more pedagogical agents, as described by Richards & Dignum in this special issue. Some evaluative studies of the systems' learning outcomes are also included in this paper. Sometimes the results were not as positive as the overall meta-analyses described above have found. For the purpose of illustrating the breadth of ILE pedagogies, this is not an issue. For a more systematic review of adaptive educational technology, see Aleven, McLaughlin, Glenn, and Koedinger (2017).

This paper is organised in a further four sections. In order to illustrate the breadth of the pedagogies now available, the first part distinguishes the *screen-level pedagogy* of how individual systems work with a single student, from the *orchestration-level pedagogy* whereby such systems are deployed in the bigger temporal and spatial context of a whole class. "Orchestration" covers a range of issues including tutoring more than one student at a time, and the integration of the system into other classroom activities.

The next two sections look at examples of screen-level and orchestration-level pedagogies, chosen to illustrate the wide range of pedagogies and deployment methods now in use. One thread that links these examples is the extent to which the different pedagogies promote *active* learning, rather than the passive learning associated with Skinner (see for example, Skinner, 1968). As the main focus is on the broad nature of the pedagogies of these systems, there is no attempt to explain how these pedagogies are achieved technically. Finally, the fourth section indicates some directions in which pedagogy might develop.

Screen-level and Orchestration-level

Researchers distinguish two levels of pedagogy for ILEs operating in formal and semi-formal education. The first is the *screen-level* pedagogy, determining the interaction of the learner with the system. Issues here include who "drives" the interaction (the learner or the system), who chooses what tasks to address (the learner or the system), what kinds of help and feedback are offered (or provided) and when, and in what ways the system adapts these and other factors to the individual learner or group of learners.

Adapting to the learner is a central feature of this screen-level pedagogy. This feature is based, in part, on the finding that the most effective tutoring is provided one-to-one by a skilled-human tutor who tailors the interaction to the specific capabilities, progress and needs of the tutee in order to achieve mastery at each stage (B. S. Bloom, 1984). Adaptation can be achieved along different learner dimensions such as their prior knowledge and skill, their knowledge and skill growth, their path through the material, their learning strategies and metacognition, their effort and their learning style (for a detailed discussion, including evidence of effectiveness of each of these kinds of adaptivity, see Aleven et al., 2017).

There is also interest in the "non-cognitive" aspects of adaptivity, where in addition to adapting to the cognitive state of the learner, ILEs take account of, and adapt to, the learner's affective and motivational state (for a review, see e.g., du Boulay, 2018). This kind of adaptivity helps with student engagement by keeping the learner interested in what is being learned and able to deal with the inevitable setbacks caused by confusion, frustration, anxiety and boredom that are commonplace experiences of learners. For an example of such a successful, multi-faceted pedagogy, see Arroyo et al. (2014). They developed pedagogic tactics to deal with a range of inferred student states, such as achieving mastery but with high effort. For each student state they devised a cognitive decision, such as to maintain problem difficulty, as well as separate affective and motivational decisions, such as to praise the effort that the student had shown.

The second level of pedagogy is the orchestration of the ILE. This addresses the issue of how the learners' work with the system(s) is organised within the wider context of other learners, of their other activities and of the course they are undertaking, i.e. in formal educational settings (Dillenbourg, 2013). Dillenbourg emphasises the fact that the human teacher has a much richer set of possible actions and adaptations available than can be designed into an ILE, and so can also exploit unexpected classroom events to enrich the learner's activities with that ILE. Moreover, the teacher is likely to have a better sense than the system of the learner as a person and of their overall personal context.

Various "orchestration systems" have been developed to help the teacher manage the competing demands on their time to assist individuals and subgroups making use of educational technology within a class (see for example, Martinez-Maldonado, Dimitriadis, Kay, Yacef, & Edbauer, 2013; Prieto, Rodríguez-Triana, Martínez-Maldonado, Dimitriadis, & Gašević, 2019). One of the primary aims of such systems is to provide the teacher with an indication of who, or which subgroup, needs their attention most at a given moment. Reasoning about these competing needs involves some of the same student modelling techniques that also underpin ILEs. For example, an orchestration system, such as FACT, changes the classroom dynamics and augments the pedagogy as it offers an extra private channel of feedback and communication between each learner and the teacher, as well as between the learners themselves (Cheema, VanLehn, Burkhardt, Pead, & Schoenfeld, 2016).

A recent study has illustrated the importance of the teacher *complementing* the scaffolding available in the system with further scaffolding that better respects the needs of the individual learner, rather than simply extending the kind of scaffolding already provided by the system (Martin, Tissenbaum, Gnesdilow, & Puntambekar, 2018). Note the finding from (Ma et al., 2014), described in the Introduction, that "Significant, positive mean effect sizes were found regardless of whether the ILE was *used as the principal means of instruction, a supplement to teacher-led instruction, an integral component of teacher-led instruction* . . .[my emphasis]."

We turn now to several exemplars of screen-level and orchestration-level pedagogy.

Screen-level pedagogy: Example Systems

This section describes a number of ILEs. There is no attempt to cover the whole field, so examples have been chosen to give a broad sense of what has been developed. The main aim is to demonstrate that these systems offer educational interactions that go far beyond "the Skinner Box". Each subsection focuses on a particular pedagogy.

Worked Examples vs. Problems

The Cognitive Tutors are probably the most well-known form of ILE for teaching skills (Anderson, Corbett, Koedinger, & Pelletier, 1995). In their classic form they are used to teach problem-solving, based on an extensive empirical investigation by the developers to determine the skills and subskills involved (e.g. algebra, geometry, statistics, chemistry, physics, medicine, programming). The systems present a sequence of problems, adjusted to the progress of the particular learner. They ensure mastery learning by ensuring that sub-skills are mastered before main skills, and provide detailed feedback on every step taken by the learner (Block & Burns, 1976; B. S. Bloom, 1968, 1971). These systems have concentrated on STEM subjects because it is relatively straightforward to build student and domain models; later sections give examples of ILEs in non-STEM areas.

Many ILEs present an adaptive sequence of problem-statements and then help the learner to solve them, based on an AI-based analysis of what ought to be the next most useful problem for that particular student to solve (Anderson et al., 1995). A different approach is to present a problem-statement and an example-solution to that problem, and then have the learner work through the solution to gain an understanding of how and why it is a solution to the given problem-statement. Again AI-based analyses are used to identify which problem statement and example-solution to present next. This is called learning with worked-examples. A variant on this is learning via *erroneous* worked-examples. Here the learner is expected to identify, fix and explain both the error and the correction for it in the erroneous example.

In one experiment, using an ILE which compared problem-solving, worked-examples and erroneous examples in the domain of chemistry, learning gains were identical across the three conditions, but the time taken to learn was fastest with the worked examples and required the least mental effort (McLaren, van Gog, Ganoe, Karabinos, & Yaron, 2016). In another experiment that compared problem-solving with learner-explanation of erroneous examples for decimals, there was no difference in terms of immediate post-test results, but a significant difference in terms of a delayed post-test in favour of the erroneous examples group (Adams et al., 2014). These researchers also found that the problem-solving group reported greater satisfaction with their interaction than the erroneous examples group, despite it being less effective. Such a result clearly has interesting implications for a pedagogy that adapts to learner motivation and affect. In a related experiment, a fixed alternating sequence of problem-solving and worked-examples for an ILE teaching the database language SQL was compared with an adaptive sequence in which the system chose whether to offer a problem, a worked-example or a faded-example. In a faded-example, one or two steps are omitted and the learner is expected to complete them. Learners using the adaptive sequence learned both more and faster than the group using the fixed sequence (Najar, Mitrovic, & McLaren, 2016).

Learning by Teaching

Betty's Brain is a system used in formal educational settings to help learners build up a conceptual understanding of a domain rather than develop a skill as in the algebra tutor mentioned above (Leelawong & Biswas, 2008). The core pedagogy is learning by teaching.

This involves either two learners taking it in turns to teach each other, or a learner and a system where the learner teaches the system (see for example, Chou & Chan, 2016; Matsuda, Cohen, & Koedinger, 2015; Palincsar & Brown, 1984; Pareto, 2014). The system is called Betty's Brain because the narrative supporting the interaction is that the learner is building up the conceptual understanding of the character in the system (Betty) and the concept map being built is thus the contents of her brain. This subterfuge aims to reduce the learner's potential affective discomfort when her creation fails the test: it is not she who failed but Betty.

The learners are offered a range of multimedia descriptive materials about a small eco-system and are expected to formalise their understanding of it by building up a concept map of the various processes that take place within that eco-system, such as the creation of carbon dioxide. The learner draws the map as a network by selecting named nodes and links. The learner can use the network to check Betty's understanding of the biological processes by reasoning via chaining along links from one node to another. She can also ask the system to make the same kind of check, as the concept map is a "shared resource" between the learner and the system. The system itself can use an AI-based analysis to check the overall accuracy of the concept map by submitting it to a range of test questions (quizzes) and seeing how well it answers them.

The system has a pedagogical agent (see Richards & Dignum, this special issue) in the form of a teacher who not only sets and marks the tests about the eco-system and can offer advice about how to proceed but, importantly, monitors the learner's use of the multimedia materials and of the concept map testing facility. This teacher can then offer metacognitive advice about the way the learner is going about her concept learning, such as not making good use of the online materials or retesting the concept map without changing it. While the overall task is set by the system, the order and manner for building the concept map is the learner's responsibility. This is an active rather than a passive learning process, and the process itself, as well as its outcome, is monitored and scaffolded.

In one evaluation three groups of fifth grade school children were compared. One group (SRL) used the system as described above, involving both metacognitive prompts to promote self-regulated learning, as well as learning by teaching. A second group (LBT) undertook learning by teaching but without metacognitive prompts. The third group (ITS) were taught the concepts directly, did not carry out learning by teaching and did not receive metacognitive prompts. The authors summarised the results as follows:

> "The directed feedback in the ITS and LBT systems enhanced student learning performance in the early sessions of the main study. This advantage disappeared in later sessions as the students in the SRL group learnt to use strategies demonstrated by the guided metacognitive feedback in their learning and self-monitoring tasks. In the transfer test, the benefits of the guided learning became even more visible as the SRL students were able to transfer their metacognitive strategies to a new learning environment, where most of the scaffolds to aid the learning task were removed." (Leelawong & Biswas, 2008, page 205).

However, it was also found that some students lacked some of the pre-requisite skills and their wider rationale, e.g. about reasoning along causal chains, so as to learn about the specific causal relationships in the eco-system to be modelled (Biswas, Segedy, & Bunchongchit, 2016). In order to improve the system's effectiveness, the introductory narrative was augmented to explain the nature of this reasoning and to make it clearer that this kind of reasoning would have many other applications.

Learning through Dialogue

Many of the examples of systems so far cited have been designed to tutor relatively closed rather than more open-ended tasks, and the communication between learner and system has not involved natural language. By contrast, ReportTutor aimed to help trainee pathologists learn how to write a diagnostic pathology report following a biopsy for melanoma (Saadawi et al., 2008). It presented images of pathology slides and then invited the learner to write a report on what the slides revealed, and possible diagnoses. This report was to be written in free prose. The learner reports were compared against a reference standard report for each case produced by an expert Dermatopathologist who “indicated the features, attribute names, and attribute values that ReportTutor is designed to detect”. The system read the report and offered comments and criticisms of what had been written, also in free prose. The roles of AI in this system were to understand the text that the learner had written and then identify ways in which the learner’s diagnostic report was either inaccurate or missing features. In terms of pedagogy, ReportTutor could operate in two modes. In one the feedback from the system was provided once the draft report was finished; in the other mode it could provide feedback immediately while the report was being written. In a comparative experiment it was found that both modes of feedback produced similar improvements in pathological report writing.

In a 17 year research program based at University of Memphis, a range of dialogue-based tutors have been developed under the generic title, AutoTutor (Nye, Graesser, & Hu, 2014). AutoTutor typically engages with the learner by questioning them about a topic, e.g. computer literacy or critical thinking, and then analysing and reacting to the adequacy of the learner’s free-form answer in English. The researchers express the pedagogical argument for this approach as follows:

> “Discourse with the student opens up a range of new learning activities related to expressing and communicating knowledge (e.g., self-reflection, answering deep questions, generating questions, resolving conflicting statements). The strengths of natural language tutoring complement a wide range of domains, including traditional problem-solving . . . . these activities emphasize comprehending, analyzing, synthesizing, and evaluating knowledge . . . . Dialog-based tutoring helps integrate concepts and problem-solving with domain principles, scaffold domain-specific language, implement abstract strategies, and generate qualitative inferences.” (Nye et al., 2014, page 430)

The quote above emphasises their interest in active modes of learning. The design of AutoTutor is based on extensive empirical research into student-teacher dialogues, including those that involve the co-creation of explanations and understanding. As a result, this research enabled the designers to incorporate a rich AI-based repertoire of conversational tutorial gambits designed to detect the learner’s state of understanding and develop it as much as possible. In addition, some versions of the system detect and respond to the learner’s motivational and affective state as well as to the quality of their answers (see, e.g. D'Mello et al., 2008). A later evaluation of the tutor found that “An experiment comparing the affect-sensitive and non-affective tutors indicated that the affective tutor improved learning for low-domain knowledge students, particularly at deeper levels of comprehension.” (D’Mello et al., 2011).

An earlier subsection described systems that made use of erroneous or faded worked-examples. A related technique in concept learning involves presenting the learner with contradictory material in order to induce confusion, followed, one hopes, by the impulse to investigate and resolve the contradiction. This was a technique adopted by Lehman et al. (2013) in a system involving more than one pedagogical agent, which provided contradictory information to the student. They found that the method worked and produced better learning than in a no-confusion control condition.

Learning about Learning

One of the major goals of education is to help learners learn better how to learn. One facet of this major goal is helping learners develop their own metacognitive monitoring and self-regulated learning (SRL) skills. The systems we have already described have tackled this issue to some extent. So, for example, in Betty's Brain, the teacher pedagogical agent monitors and advises the learner about the way they are going about learning, and in the ILEs for mathematics or programming, the use of faded or erroneous worked examples requires learners to generate explanations for the nature of missing or incorrect steps.

Azevedo coined the term "computers as *Meta*Cognitive tools" [my emphasis] as a natural development of educational technology systems that act simply as cognitive tools (Azevedo, Witherspoon, Chauncey, Burkett, & Fike, 2009). One characteristic of such a tool is that by using AI, "it models, prompts, and supports learners (to some degree) to engage or participate (alone, with a peer, or within a group) in using task-, domain-, or activity-specific learning skills". Azevedo et al. (2009) describe MetaTutor, a non-adaptive hypermedia system for learning about body systems such as the nervous system. One of their evaluations showed that typical high school and college level students "have little declarative knowledge of key SRL processes" and learned significantly more when assisted by MetaTutor.

A later, adaptive version of MetaTutor was used to develop and evaluate the note-taking behaviour of college level students learning about the human circulatory system (Trevors, Duffy, & Azevedo, 2014). The goal was to tackle a particular aspect of self-regulation in learning, namely, to reduce the extent to which learners undertook shallow note taking: essentially just reproducing the source material, rather than elaborating it. The system had four AI-based pedagogical agents – a guide, a planner, a monitor and a strategiser – that were designed to assist with different aspects of the learners' monitoring and regulation of their learning. These agents offered hints to learners in the experimental condition of the evaluation, but were silent in the control condition. The results were mixed, indicating the complexity of the interactions between learning, self-regulation of learning, and hints and advice about the latter. For example, learners with low prior knowledge of the circulatory system produced fewer notes as a result of the advice of the system, but these notes did not show evidence of high-level elaboration. Learners with high prior knowledge did not change their note taking behaviour.

Deployment-level Pedagogy: Example Systems

This section describes a range of pedagogies that go beyond simply a single learner working at a single screen. As in the previous section, there is no attempt to provide a comprehensive review. The examples are again chosen to offer a sense of how far beyond the Skinner Box paradigm ILEs have developed. The following subsections cover different methods of orchestration such as working in pairs and flipped learning.

The purveyors of the Cognitive Tutors, mentioned earlier, take a lot of trouble to educate teachers about how to make the best use of their intelligent tutoring systems as part of their classroom teaching (Koedinger & Aleven, 2016; Koedinger, Anderson, Hadley, & Mark, 1997). In other words, the orchestration-level is carefully managed. This means helping teachers think about which learners in a class might, in principle, benefit from the extra one-to-one adaptive tuition and about how to prepare those learners to work with the systems. It also means helping the teacher think about the effect on those learners *not* using the system.

A multi-school USA-based evaluation of the Algebra Cognitive Tutor compared learning outcomes in matched schools (both middle and high) where the experimental schools deployed the tutor as they saw fit within the context of their existing mathematics teaching (Pane, Griffin, McCaffrey, & Karam, 2014). Importantly, the careful orchestration-level pedagogy described above was not offered to the experimental schools. The control schools carried on teaching in whatever way they normally did. The evaluation was carried out over two years and produced mixed evaluation results, but with the experimental high schools producing better learning outcomes than the control schools in the second year. The fact that the results in the second year of the evaluation were better than in the first year implies that the teachers in the experimental high schools had started to improve their own orchestration-level pedagogy for this system. This further emphasises the importance of considering the orchestration-level pedagogy in evaluating ILEs.

Working in Pairs

Tutors such as the Cognitive Tutor for Algebra do not have to be used one-to-one with a single student. One possible orchestration-level strategy is to assign the tutor to work with pairs or groups of students. For example, Walker, Rummel, and Koedinger (2009) explored how the algebra tutor could be augmented to work with pairs of students engaging in peer tutoring. In this case the tutoring system had the dual role of both teaching algebra as well as supporting the mutual peer tutoring activity of the two students. The experiment compared three groups. One group involved pairs of students who received algebra tuition and AI-based *adaptive* support for their peer tutoring activity. Another group involved pairs of students who received algebra tuition and *fixed* support for their peer tutoring activity. The third group consisted of students who received individual algebra tutoring only. There were no significant differences in learning outcomes, but there were strong differences in learning behaviours: for example, the number of problems solved. A similar experiment, with similar results, was performed by Harsley, Di Eugenio, Green, and Fossati (2017) using an ILE for programming, where the pairs of students collaborated rather than engaging in peer tutoring. There were two conditions: one where there was no feedback on the quality of the pair's collaboration and another where there was. The issue for the current paper is not whether working in pairs was better or worse than working individually, but the fact that ILEs can be orchestrated to work with and support pairs of learners.

Flipped Learning

As a further orchestration pedagogy, ILEs can be used to help with homework rather than as classroom assistants. The Andes system for undergraduate physics was used in such a way (VanLehn et al., 2005). Like the Cognitive Tutors it analysed each step that the learner made in solving a multi-step physics problem and provided immediate feedback and hints. Using AI, it also engaged learners in a dialogue both about the approach to solving a particular problem, i.e. which physics principle should be used to solve the problem, as well as helping them to conceptualise the problem. The authors of the system evaluated it each academic year over 5 years, comparing cohorts of students who used the Andes system to help with their homework compared to those that did not. There were consistently significantly better scores in the exam results for those who used Andes every year with an overall effect size of 0.61. Clearly there may be some selection bias in such a result, but this does not undermine the fact of this use of an ILE in a flipped fashion.

Classroom Orchestration Systems
In terms of maximising the effectiveness of the scarce resource of the teacher's time and attention, training and support is also needed to help the teacher decide when to intervene and mediate the interaction between a learner, or a group of learners, and the intelligent tutoring system. Support is also needed to help the teacher ensure that the learners reflect effectively on their experience of the interaction with the system, e.g. via conversations both with other learners who have also used the systems and potentially also with those who have not. So, the implicit educational ethos is that the combination of the screen level pedagogy of the system together with orchestration-level pedagogy of the teacher offers an effective learning experience and guards against learner passivity and other mal-adaptive learner behaviours. The specific details of the orchestration of a system comes into even sharper focus when the educational, cultural context in which it is to be used differs from that in which it was developed (Casas, Fernandez, Barrera, & Ogan, 2015).

Holstein, McLaren, and Aleven (2018) have developed an augmented reality orchestration system, Lumilo, for a teacher managing a class of students using an intelligent tutoring system for mathematics. This is similar to the FACT system mentioned earlier (Cheema et al., 2016), based on tablet technology, but in this case the teacher's system provides "spectacles" that allow the teacher to observe an augmented reality view of a class of students using the intelligent tutoring system. The AI-based augmented view includes an overall listing of the common difficulties across the whole group, as well as virtual symbols hovering above each learner indicating individual levels of progress and non-progress as well as requests for help. These symbols include idle, rapid attempts, hint abuse or gaming the system, high error after hints, hint avoidance, and unproductive persistence. The system is designed to help the teacher make best use of her time to identify which issues might need more of her input when she is teaching the whole class, as well as identifying which learners might need more immediate, individual help than the system is able to provide on its own.

Developing New Pedagogy
This paper has argued that one needs to consider both the pedagogy of the system itself, as well as the pedagogy of the way the teacher makes use of the system. Indeed, one of the future directions of the development of pedagogy involves extending the capability of orchestration systems such as FACT and Lumilo. These expansions might include, for example, giving the teacher more scope to choose which learner behaviours the system flags up as needing their intervention as the teacher, assisting the teacher in forming subgroups within a class with a view to fostering collaborative learning, each group with their most appropriate tasks to work on, or pairing up and assisting students to engage in peer tutoring. The design of adaptive systems is now being augmented to take account of, and adapt the social relationships between learners, between groups and between learners and teachers (Walker & Ogan, 2016).

The idea that the pedagogy of orchestration is as important as the pedagogy of the systems being orchestrated is not always understood. For example, the guidance on implementing adaptive courseware from the American Association of Public and Land-grant Universities does not mention the issue (see https://www.aplu.org/news-and-media/News/aplu-releases-first-of-its-kind-guide-for-implementing-adaptive-courseware).

One of the weaknesses of ILEs is that they have little understanding of the context within which they are being used, and although they track the learner's actions with the system in great detail, they do not really know much about the learners as people, in the way that a human teacher usually does. So, if an intelligent learning environment for algebra (say) or ecosystems (say) is introduced into a classroom, it is currently the teacher's job as part of orchestration to contextualise the experience, skills and concepts that the learners will encounter.

It is also the teacher's job to make the case as to why it is worth the learner putting some effort into the interaction, rather than just idly pressing the buttons, misusing the help, or gaming the system (R. Baker et al., 2008). This goes beyond simply dealing with issues around the learner misunderstanding the relation between effort and learning and the importance of resilience (Yeager & Dweck, 2012) or developing self-regulation capability (Azevedo & Aleven, 2013), but focuses on what the *value* of this learning might be for *that* learner which might make the effort worthwhile.

While progress has been made in developing an affective and motivational pedagogy (for a review, see du Boulay, 2018), the main pedagogical strategy is still largely *reactive* to the learner's affective and motivational state by providing tactics to deal with difficulties as they arise, rather than *proactive*, anticipating and ameliorating affective and motivational problems before they occur (see for example, Rosiek, 2003).

We can anticipate that one line of development of ILEs would involve paying more intention to understanding and scaffolding the learner's initial volition – the will to learn (Keller, 2008) – and to the longer-term goal of much education, which is to inculcate a love of learning in general (Maehr, 2012). To achieve this will require systems to be able to adapt the manner in which they represent what the learner is about to experience to suit different learners. They will also need to assist the learner, at various points to understand the relevance of their learning effort to their longer-term goals (see, for example, the "Relevance Enhancing Strategies" in Song & Keller, 2001). Finally, they will need to help the learner engage in a post-learning reflective analysis of "how it went" for them and how future learning might be affected.

Stimulating Initial Interest

An example of stimulating initial interest comes from Coach Mike, a system deployed in a museum setting rather than a classroom. Coach Mike is a pedagogical agent deployed in a science and technology museum whose aim is to entice visitors to pay attention to, and then engage with, an interactive exhibit about programming (Lane et al., 2013). We can regard it as a system playing a role in a semi-formal educational environment. Such systems:

> ". . . go beyond simply focuses [sic] on knowledge outcomes. They must take seriously goals such as convincing a visitor to engage, promoting curiosity and interest, and ensuring that a visitor has a positive learning experience. In other words, pedagogical agents for informal learning need to not only act as coach (or teacher), but also as advocate (or salesperson)" (Lane et al., 2013, page 310).

This system operates in two modes. First it is a “salesman” to persuade a visitor to pay some attention to the interactive exhibit. Then it is a “cheerleader” as it enthusiastically helps the visitor learn to program a moveable robot, using physical blocks where each block corresponds to a command, such as LEFT, FORWARD, and SPIN. In such an environment the visitor can move away from the exhibit at any time, so one of the goals of Coach Mike is persuade the visit to stay and remain interested.

Adjusting the Way Things are to be Learned
Although there are several examples of language based tutors (see e.g., Nye et al., 2014), we are a long way from systems being able to have natural language dialogues with learners about their *interest* in a topic and their *potential willingness to put effort* in that day. However, a “personalisation” menu-based interaction together with data about other students’ preferences might establish the values of a number of parameters that could be used to recommend which learning objects to use: see for example systems such as Century (www.century.tech) and Sana Labs (www.sanalabs.com). There is still much scope for systems to adjust the way that the topic is introduced, to provide compelling examples of the relationships of the topic to other things that the learner is interested in, and to manage the order and manner in which different parts of the topic are dealt with. These examples might be repeated in some form when the system detects that the learner’s interest is flagging, e.g. as evidenced by gaming the system, abusing the help system or not making much progress.

An early example of the above was a study advisor, SOLA* (Arshad & Kelleher, 1993). This system was aimed at students who were proposing to engage in self-directed study in statistics, but it could be adapted to other domains and other learning situations. The system conducted an initial interaction with the student to determine their “background knowledge, the relevance of the course to their degree and the time which they have available in which to learn”. The system also ascertained “their levels of confidence, degree of support required (e.g. directed, exploratory) and whether they want passive/active learning methods.”

R. S. Baker (2016) argues that the cost and time needed to develop ILEs has meant that the most widely-adopted systems have not incorporated some of the best features developed and evaluated in research labs. He also argues that there is a parallel and necessary development in the use of data analytics to develop pedagogy. He points to such systems as Course Signals (Arnold & Pistilli, 2012) that assist human teachers manage whole cohorts of students, e.g. to spot potential drop outs, as opposed to whole classes of students managed by orchestration systems. He also praises systems such as ASSISTments, mentioned earlier, for their capability of running educational experiments and thus collecting valuable data to assist pedagogic development (Heffernan & Heffernan, 2014).

Conclusion
This paper has challenged the comment in the call for papers in this special issue that “. . . analytics and AI equate to adopting a retrograde pedagogy from the industrial era. . .” by exploring the ongoing developments in pedagogy embodied in ILEs and anticipated some future directions. It has distinguished between screen-level pedagogy and orchestration-level pedagogy as each contribute to the breadth of the learning experience. The paper has argued that ILEs have embraced many ways to support active learning at screen-level as well as via careful orchestration of screen-level work in the wider context of the classroom.

The paper has sketched some possible future directions involving taking better account of fostering the learner's volition and making more dramatic efforts to present and organise the material to be learned in a way that is more adaptive to the interests of the learner. It has also outlined possibilities for supporting the teacher better in terms of managing the many demands on their time when learners are using ITSs and ILEs as well as other resources.

Acknowledgements

I thank members of the Human Centred Technology research group at the University of Sussex for their very helpful comments on an earlier draft of this paper and the Editor and reviewers of this paper for their insightful critiques.

Statements

I have no conflict of interest arising from writing this paper. The paper does not describe an empirical study that I conducted so ethical approval was not needed and no participant data was generated.

References

Adams, D. M., McLaren, B. M., Durkin, K., Mayer, R. E., Rittle-Johnson, B., Isotani, S., & van Velsen, M. (2014). Using erroneous examples to improve mathematics learning with a web-based tutoring system. *Computers in Human Behavior, 36*, 401-411.

Aleven, V., McLaughlin, E. A., Glenn, R. A., & Koedinger, K. R. (2017). Instruction Based On Adaptive Learning Technologies. In R. E. Mayer & P. A. Alexander (Eds.), *Handbook of Research on Learning and Instruction*: Routledge.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive Tutors: Lessons Learned. *The Journal of the Learning Sciences, 4*(2), 167-207.

Arnold, K. E., & Pistilli, M. D. (2012). *Course Signals at Purdue: Using Learning Analytics to Increase Student Success*. Paper presented at the LAK '12 Proceedings of the 2nd International Conference on Learning Analytics and Knowledge Vancouver, Canada.

Arroyo, I., Woolf, B. P., Burleson, 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*(4), 387-426.

Arshad, F. N., & Kelleher, G. (1993). SOLA*: Students On-Line Advisor. *International Journal of Man-Machine Studies, 38*(2), 281-312. doi:https://doi.org/10.1006/imms.1993.1013

Azevedo, R., & Aleven, V. (Eds.). (2013). *International Handbook of Metacognition and Learning Technologies*: Springer.

Azevedo, R., Witherspoon, A., Chauncey, A., Burkett, C., & Fike, A. (2009). *MetaTutor: A MetaCognitive Tool for Enhancing Self-Regulated Learning*. Paper presented at the AAAI Fall Symposium (FS-09-02).

Baker, R., Walonoski, J., Heffernan, N., Roll, I., Corbett, A., & Koedinger, K. (2008). Why Students Engage in "Gaming the System" Behaviours in Interactive Learning Environments. *Journal of Interactive Learning Research, 19*(2), 185-224.

Baker, R. S. (2016). Stupid Tutoring Systems, Intelligent Humans. *International Journal of Artificial Intelligence in Education, 26*(2), 600-614. doi:http://dx.doi.org/10.1007/s40593-016-0105-0

Biswas, G., Segedy, J. R., & Bunchongchit, K. (2016). From Design to Implementation to Practice a Learning by Teaching System: Betty's Brain. *International Journal of Artificial Intelligence in Education, 26*(1), 350-364. doi:http://dx.doi.org/10.1007/s40593-015-0057-9

Block, J. H., & Burns, R. B. (1976). Mastery Learning. *Review of Research in Education, 4*, 3-49.

Bloom, B. S. (1968). *Learning for Mastery*. Retrieved from

Bloom, B. S. (1971). Mastery learning. In J. H. Block (Ed.), *Mastery learning, theory and practice* (pp. 47-63). New York: Holt, Rinehart, and Winston.

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*(6), 4-16.

Bloom, C. P. (1995). *Roadblocks to successful ITS authoring in industry*. Paper presented at the AIED'95 Workshop on Authoring Shells for Intelligent Tutoring Systems, Washington, DC, USA.

Burton, R. R., & Brown, J. S. (1979). An Invetsigation of Computer for Informal Learning Activities. *International Journal Man-Machine Studies, 11*(1), 5-24.

Casas, I., Fernandez, P., Barrera, M., & Ogan, A. (2015). *Culture-Oriented Factors in the Implementation of Intelligent Tutoring Systems in Chile*. Paper presented at the Sixth International Workshop on Culturally-Aware Tutoring Systems (CATS) at AIED 2015, Madrid, Spain.

Cheema, S., VanLehn, K., Burkhardt, H., Pead, D., & Schoenfeld, A. (2016). *Electronic Posters to Support Formative Assessment*. Paper presented at the CHI EA '16 Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems San Jose, CA, USA.

Chou, C.-Y., & Chan, T.-W. (2016). Reciprocal Tutoring: Design with Cognitive Load Sharing. *International Journal of Artificial Intelligence in Education, 26*(1), 512-535. doi:http://doi.org/10.1007/s40593-015-0083-7

Clancey, W. J. (1979). Tutoring Rules for Guiding a Case mrthod Dialogue. *International Journal Man-Machine Studies, 11*(1), 25-50.

D'Mello, S., Jackson, T., Craig, S., Morgan, B., Chipman, P., White, H., . . . Picard, R. (2008). *AutoTutor Detects and Responds to Learners Affective and Cognitive States*. Paper presented at the Workshop on Emotional and Cognitive Issues at the International Conference on Intelligent Tutoring Systems, Montreal, Canada.

D'Mello, S., Lehman, B., Sullins, J., Daigle, R., Combs, R., Vogt, K., . . . Graesser, A. (2011). A Time For Emoting: When Affect-Sensitivity Is and Isn't Effective at Promoting Deep Learning. In R. A. Calvo & S. K. D'Mello (Eds.), *Affective Prospecting* New York: Springer.

Dillenbourg, P. (2013). Design for classroom orchestration. *Computers & Education, 69*, 485-492. doi:https://doi.org/10.1016/j.compedu.2013.04.013

du Boulay, B. (2016). Artificial Intelligence As An Effective Classroom Assistant. *IEEE Intelligent Systems, 31*(6), 76-81.

du Boulay, B. (2018). Intelligent Tutoring Systems That Adapt to Learner Motivation. In S. D. Craig (Ed.), *Tutoring and Intelligent Tutoring Systems* (pp. 103-128). New York: Nova Science Publishers, Inc.

Harsley, R., Di Eugenio, B., Green, N., & Fossati, D. (2017). *Collaborative Intelligent Tutoring Systems: Comparing Learner Outcomes Across Varying Collaboration Feedback Strategies*. Paper presented at the 12th International Conference on Computer Supported Collaborative Learning (CSCL 2017), Philadelphia, Pennsylvania, USA.

Heffernan, N. T., & Heffernan, C. L. (2014). The ASSISTments ecosystem: building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. *International Journal of Artificial Intelligence in Education, 24*(4), 470-497. doi:https://dx.doi.org/10.1007/s40593-014-0024-x

Herold, B. (2017). The Case(s) Against Personalized Learning. *Education Week, 37,* 4-5.

Holstein, K., McLaren, B. M., & Aleven, V. (2018). Student Learning Benefits of a Mixed-Reality Teacher Awareness Tool in AI-Enhanced Classrooms. In C. P. Rosé, R. Martínez-Maldonado, H. U. Hoppe, R. Luckin, M. Mavrikis, K. Porayska-Pomsta, B. McLaren, & B. du Boulay (Eds.), *Artificial Intelligence in Education: 19th International Conference, AIED 2018, London, UK, June 27–30, 2018 Proceedings, Part I* (pp. 154-168): Springer.

Keller, J. M. (2008). An Integrative Theory of Motivation,Volition, and Performance. *Technology, Inststruction, Cognition and Learning, 6*(2), 79-104.

Koedinger, K. R., & Aleven, V. (2016). An Interview Reflection on "Intelligent Tutoring Goes to School in the Big City". *International Journal of Artificial Intelligence in Education, 16*(1), 13-24. doi:http://dx.doi.org/10.1007/s40593-015-0082-8

Koedinger, K. R., Anderson, J. R., Hadley, W. H., & Mark, M. A. (1997). Intelligent Tutoring Goes to School in the Big City. *International Journal of Artificial Intelligence in Education, 8*(1), 30-43.

Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of Intelligent Tutoring Systems: A Meta-Analytic Review. *Review of Educational Research, 86*(1), 42-78. doi:http://dx.doi.org/10.3102/0034654315581420

Lane, H. C., Cahill, C., Foutz, S., Auerbach, D., Noren, D., Lussenhop, C., & Swartout1, W. (2013). The Effects of a Pedagogical Agent for Informal Science Education on Learner Behaviors and Self-efficacy. In *Artificial Intelligence in Education: 16th International Conference, AIED 2013, Memphis, TN, USA, July 9-13, 2013. Proceedings* (pp. 309-318). Berlin: Springer.

Leelawong, K., & Biswas, G. (2008). Designing Learning by Teaching Agents: The Betty's Brain System. *International Journal of Artificial Intelligence in Education, 18*(3), 181-208.

Lehman, B., D'Mello, S., Strain, A., Mills, C., Gross, M., Dobbins, A., . . . Graesser, A. C. (2013). Inducing and Tracking Confusion with Contradictions during Complex Learning. *International Journal of Artificial Intelligence in Education, 22*(1-2), 85-105.

Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent Tutoring Systems and Learning Outcomes: A Meta-Analysis. *Journal of educational psychology, 106*(4), 901-918. doi:http://dx.doi.org/10.1037/a0037123

Maehr, M. L. (2012). *Encouraging a Continuing Personal Investment in Learning: Motivation As an Instructional Outcome*. Charlotte, NC, USA: Information Age Publishing.

Martin, N. D., Tissenbaum, C. D., Gnesdilow, D., & Puntambekar, S. (2018). Fading distributed scaffolds: the importance of complementarity between teacher and material scaffolds. *Instructional Science*, 1-30. doi:https://doi.org/10.1007/s11251-018-9474-0

Martinez-Maldonado, R., Dimitriadis, Y., Kay, J., Yacef, K., & Edbauer, M.-T. (2013). *MTClassroom and MTDashboard: Supporting analysis of teacher attention in an orchestrated multi-tabletop classroom*. Paper presented at the International Conference on Computer Supported Collaborative Learning (CSCL2013), Madison, Wisconsin, USA.

Matsuda, N., Cohen, W. W., & Koedinger, K. R. (2015). Teaching the Teacher: Tutoring SimStudent Leads to More Effective Cognitive Tutor Authoring. *International Journal of Artificial Intelligence in Education, 25*(1), 1-34. doi:https://dx.doi.org/10.1007/s40593-014-0020-1
McLaren, B. M., van Gog, T., Ganoe, C., Karabinos, M., & Yaron, D. (2016). The efficiency of worked examples compared to erroneous examples, tutored problem solving, and problem solving in computer-based learning environments. *Computers in Human Behavior, 55, Part A*, 87-99.
Najar, A. S., Mitrovic, A., & McLaren, B. M. (2016). Learning with intelligent tutors and worked examples: selecting learning activities adaptively leads to better learning outcomes than a fixed curriculum. *User Modeling and User-Adapted Interaction, 26*(5), 459-491.
Nesbit, J. C., Adesope, O. O., Liu, Q., & Ma, W. (2014). *How Effective are Intelligent Tutoring Systems in Computer Science Education?* Paper presented at the IEEE 14th International Conference on Advanced Learning Technologies (ICALT), Athens, Greece.
Nye, B. D., Graesser, A. C., & Hu, X. (2014). AutoTutor and Family: A Review of 17 Years of Natural Language Tutoring. *International Journal of Artificial Intelligence in Education, 24*(4), 427-469. doi:http://dx.doi.org/10.1007/s40593-014-0029-5
O'Shea, T. (1979). A Self-Improving Quadratic Tutor. *International Journal Man-Machine Studies, 11*(1), 97-124.
Palincsar, A. S., & Brown, A. L. (1984). Reciprocal Teaching of Comprehension-Fostering and Comprehension-Monitoring Activities. *Cognition and Instruction, 1*(2), 117-175.
Pane, J. F., Griffin, B. A., McCaffrey, D. F., & Karam, R. (2014). Effectiveness of Cognitive Tutor Algebra I at Scale. *Educational Evaluation and Policy Analysis, 36*(2), 127-144. doi:http://dx.doi.org/10.3102/0162373713507480
Pareto, L. (2014). A Teachable Agent Game Engaging Primary School Children to Learn Arithmetic Concepts and Reasoning. *International Journal of Artificial Intelligence in Education, 24*(3), 251-283. doi:http://doi.org/10.1007/s40593-014-0018-8
Prieto, L. P., Rodríguez-Triana, M. J., Martínez-Maldonado, R., Dimitriadis, Y., & Gašević, D. (2019). Orchestrating Learning Analytics (Orla): Supporting Inter-Stakeholder Communication About Adoption of Learning Analytics at the Classroom Level. *Australasian Journal of Educational Technology, 35*(4), 14-29. doi:https://doi.org/10.14742/ajet.4314
Roschelle, J., Feng, M., Murphy, R. F., & Mason, C. A. (2016). Online Mathematics Homework Increases Student Achievement. *AERA Open, 2*(4), 1-12. doi:https://doi.org/10.1177/2332858416673968
Rosenberg, R. (1987). A Critical Analysis of Research on Intelligent Tutoring Systems *Educational Technology, 27*(11), 7-13.
Rosiek, J. (2003). Emotional Scaffolding: An Exploration of the Teacher Knowledge at the Intersection of Student Emotion and the Subject Matter. *Journal of Teacher Education, 54*(4), 399-412.
Saadawi, G. M. E., Tseytlin, E., Legowski, E., Jukic, D., Castine, M., Fine, J., . . . Crowley, R. S. (2008). A Natural Language Intelligent Tutoring System for Training Pathologists - Implementation and Evaluation. *Advances in Health Sciences Education, 13*(5), 709-722. doi:https://dx.doi.org/10.1007/s10459-007-9081-3
Skinner, B. F. (1968). *The Technology of Teaching*. New York: Appleton-Century-Crofts.
Song, S. H., & Keller, J. M. (2001). Effectiveness of motivationally adaptive computer-assisted instruction on the dynamic aspects of motivation. *Educational technology research and development, 49*(2), 5-22.

Steenbergen-Hu, S., & Cooper, H. (2013). A Meta-Analysis of the Effectiveness of Intelligent Tutoring Systems on K–12 Students' Mathematical Learning. *Journal of educational psychology, 105*(4), 970-987. doi:http://dx.doi.org/10.1037/a0032447

Steenbergen-Hu, S., & Cooper, H. (2014). A meta-analysis of the effectiveness of intelligent tutoring systems on college students' academic learning. *Journal of educational psychology, 106*(2), 331-347. doi:http://dx.doi.org/10.1037/a0034752

Tabor, N. (2018). Brooklyn Students Are Protesting Silicon Valley's Favorite Education Program. *Intelligencer*.

Trevors, G., Duffy, M., & Azevedo, R. (2014). Note-taking within MetaTutor: interactions between an intelligent tutoring system and prior knowledge on note-taking and learning. *Educational technology research and development, 62*(5), 507-528. doi:https://www.doi.org/10.1007/s11423-014-9343-8

VanLehn, K. (2011). The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems. *Educational psychologist, 46*(4), 197-221. doi:http://dx.doi.org/10.1080/00461520.2011.611369

VanLehn, K., Lynch, C., Schulze, K., Shapiro, J. A., Shelby, R., Taylor, L., . . . Wintersgill, M. (2005). The Andes Physics Tutoring System: Lessons Learned. *International Journal of Artificial Intelligence in Education, 15*(3), 147-204.

Walker, E., & Ogan, A. (2016). We're in this Together: Intentional Design of Social Relationships with AIED Systems. *International Journal of Artificial Intelligence in Education, 26*(2), 731-729.

Walker, E., Rummel, N., & Koedinger, K. R. (2009). Integrating Collaboration and Intelligent Tutoring Data in the Evaluation of a Reciprocal Peer Tutoring Environment. *Research and Practice in Technology Enhanced Learning, 4*(3), 221-251.

Watters, A. (2015). Education Technology and Skinner's Box. Retrieved from http://hackeducation.com/2015/02/10/skinners-box

Watters, A. (2017). Dunce's App: How Silicon Valley's brand of behaviorism has entered the classroom. *The Baffler*.

Williamson, B. (2018a). Learning lessons from data controversies. Retrieved from https://codeactsineducation.wordpress.com/2018/12/18/learning-lessons-from-data-controversies/

Williamson, B. (2018b). Silicon startup schools: technocracy, algorithmic imaginaries and venture philanthropy in corporate education reform. *Critical Studies in Education, 59*(2), 218-236. doi:https://doi.org/10.1080/17508487.2016.1186710

Williamson, B. (2018c). The tech elite is making a power-grab for public education. Retrieved from https://codeactsineducation.wordpress.com/2018/09/14/new-tech-power-elite-education/

Wilson, C., & Scott, B. (2017). Adaptive systems in education: a review and conceptual unification. *The International Journal of Information and Learning Technology, 34*(1), 2-19. doi:http://dx.doi.org/10.1108/IJILT-09-2016-0040

Yeager, D. S., & Dweck, C. S. (2012). Mindsets That Promote Resilience: When Students Believe That Personal Characteristics Can Be Developed. *Educational psychologist, 47*(4), 302-314. doi:http://dx.doi.org/10.1080/00461520.2012.722805