Workshop on Les Contes du Mariage: Should AI stay married to Ed?

held in conjunction with

Seventeenth International Conference on
Artificial Intelligence in Education (AIED 2015)

Monday, June 22, 2015
Madrid, Spain

Workshop Co-Chairs:

Kaska Porayska-Pomsta,[1] Gord McCalla[2] and
Benedict du Boulay[3]

[1]UCL Institute of Education, London, WC1H 0AL, UK
[2]University of Saskatchewan, Saskatoon, S7N 5C9, Canada,
[3]University of Sussex, Brighton, BN1 9QJ, UK

http://www.sussex.ac.uk/Users/bend/aied2015

Table of Contents

Preface i - ii

Whither or wither the AI of AIED? 1-10
Judy Kay

A is for Adaptivity, but What is Adaptivity? Re-Defining the Field of AIED 11-20
Vincent Aleven

All that glitters (in the lab) may not be gold (in the field) 21-27
Amruth N. Kumar

Why AIED Needs Marriage Counselling by Cognitive Science (to Live Happily Ever After) 28-37
Björn Sjödén

AI and Education: Celebrating 30 years of Marriage 38-45
Beverly Park Woolf

AI and Ed: a Happy Open Marriage 48-51
Julita Vassileva, James Lester, Judith Masthoff

AI in Education as a methodology for enabling educational evidence-based practice 52-61
Kaska Porayska-Pomsta

AIED Is Splitting Up (Into Services) and the Next Generation Will Be All Right 62-71
Benjamin D. Nye

Education still needs Artificial Intelligence to support Personalized Motor Skill Learning: Aikido as a case study 72-81
Olga C. Santos

Realizing the Potential of AIED 82-85
Lewis Johnson

Preface

At its origin, the field of Artificial Intelligence in Education (AIEd) aimed to employ Artificial Intelligence techniques in the design of computer systems for learning. The 25th anniversary of the IJAIEd is a good opportunity to interrogate the aims and aspirations of the field, its past and current achievements, while the AIED conference constitutes a timely forum for such an interrogation. This workshop explores questions such as:

- What is and what should be the role of AI in Education and conversely of Education in AI? Specifically, in the early days of AIED there seemed to be lots of AI in AIED, but now AI is more often a placeholder for any kind of advanced technology.
- What is and what should be the motivation of AIEd as a field? Supporting learning has been considered a great "challenge domain" for AI in that many of the big AI questions must be answered, at least to some extent, to build a sophisticated learning environment. But, it seems that the ideas generated in AIED are neither influencing AI nor Education in any serious way. Why not?
- What is and what should be the balance of respective contributions to AIED from AI and Education as distinct fields of research and practice? Both fields have well-established methodologies and practices, but the extent to which these are cross-fertilising under AIED is not clear.
- A related question relates to the extent to which the results of AIEd research are meaningful to real educational practices? Does the community even care?
- What are the future directions for the field that could justify and maintain its unique identity? How does AIED differ from related disciplines such as Learning Sciences, ITS, and CSCL? Or are these just labels for essentially the same research discipline?

We thank the many people who contributed toward making this workshop a reality. In particular, we are grateful to the authors and to the Programme Committee, listed below:

- Vincent Aleven, Carnegie Mellon University, USA
- Nicolas Balacheff, Centre National de la Recherche Scientifique, Grenoble, France
- Sara Bernardini, King's College, London, UK
- Bill Clancey, Florida Institute for Human and Machine Cognition, USA
- Tak-Wai Chan, National Central University, Taiwan
- Pierre Dillenbourg, Swiss Federal Institute of Technology, Lausanne, Switzerland
- Vania Dimitrova, University of Leeds, UK
- Claude Frasson, University of Montreal, Canada

- Jim Greer, University of Saskatchewan, Canada
- Lewis Johnson, President and CEO, Alelo Inc., USA
- Judy Kay, University of Sydney, Australia
- Chad Lane, University of Illinois, USA
- James Lester, NC State University, USA
- Chee-Kit Looi, National Institute of Education, Singapore
- Rose Luckin, UCL Institute of Education, London, UK
- Manolis Mavrikis, UCL Institute of Education, London, UK
- Richard Noss, UCL Institute of Education, London, UK
- Helen Pain, University of Edinburgh, UK
- Elliot Soloway, University of Michigan, USA
- Julita Vassileva, University of Saskatchewan, Canada
- Beverly Woolf, University of Massachusetts, USA

Kaska Porayska-Pomsta, Gord McCalla and Benedict du Boulay
May 2015

Whither or wither the AI of AIED?

Judy Kay

School of Information Technologies, University of Sydney, Australia
judy.kay@sydney.edu.au

Abstract. This position paper explores the relationship between the historic roots of AIED and the challenges of restricting our vision to EdTech that has AI. It argues that the founders of AIED had a broad vision of the field, primarily driven by the *goals* of creating advanced technology for personalised learning. They were not wedded to a techno-centric view, demanding use of particular techniques that are now thought of as "AI". The paper argues that we have accepted work with no AI, notably in Open Learner Modelling. We discourage, either directly or just because of our name, work that is true to the AIED founders' vision. In doing so, we miss many exciting and promising ways to create better technology for education.

1 What was the AI in the initial vision of AIED?

So how did we come to be called AIED in the first place? In the early days of computing research, AI had a very broad brief. It was driven by the vision that computers would one day be able to emulate the actions we describe as intelligent when people do them. What a bold vision this was --- at a time when computers were very slow, expensive and available only in research labs, military and business contexts. AI research stood in stark contrast to other the major areas of computing, such as hardware, operating systems, programming languages and numerical analysis. It was AI that looked to real world applications and creating the visions of science fiction.

AIED was born in the 1970s, with its first conferences in the 1980s (Self, 2015). It aspired to create applications that could help people learn. This was long before it was possible for most learners to even see, let alone use, a computer. A widely cited driver for our AIED research was the vision that computers could help achieve Bloom's famous 2-sigma learning benefits from *personalized teaching* by an expert teacher (Bloom, 1984). Our community is still committed to this goal. But it is useful to consider what it meant.

The classic early work in AIED identified four key elements:

- domain expertise;
- teaching expertise;
- student model; and
- user interface.

And so, the goal of researchers was to explore any or all of these architectural elements, towards building what was called an Intelligent Tutoring Systems (ITS) or AIED system. Overall, for both AIED and ITS, one key goal was to create computer systems that could provide *personalised teaching*, just as a knowledgeable teacher with expert teaching skills could do. We still aim to do this. And another goal was to support excellent *user interfaces* --- with what we may now call natural user interfaces (such as natural language and speech) and rich forms of interaction (such as graphical user interfaces that are now the norm). The spirit of their vision included creating systems and interfaces that both mimic human expert teaching and to use other techniques that are better suited to machines.

Since our early days, when the AIED community chose its name, a great deal has changed for AI, computing broadly, even for the behemoth of formal education and the commercial interests associated with those institutions and broader education. In parallel, AIED research has evolved in important ways. The next part of this paper explores these differences as a foundation for arguing that AI still has a place in AIED, but that it is not necessary for the still worthy and, as yet, unreached core vision of our founders.

2 How has AI changed since the birth and naming of AIED?

AI has become mainstream in the sense that it is part of the technology that each of us uses each day. This is well illustrated in the following descriptions from the EdX Introduction to AI[1].

Artificial intelligence is already all around you, from web search to video games. AI methods ***plan*** your driving directions, ***filter*** your spam, and ***focus your cameras on***

[1] https://www.edx.org/course/artificial-intelligence-uc-berkeleyx-cs188-1x-0#.VQzBWUaI0k4

faces. AI lets you guide your phone with your ***voice*** and ***read foreign newspapers in English***. Beyond today's applications, AI is at the core of many new technologies that will shape our future. From self-driving cars to household robots, advancements in AI help transform science fiction into real systems.

I have added the bold font to highlight the sampler of technical areas alluded to: planning, filtering, vision, natural language translation. AI has been so successful that it has resulted in many off-the-shelf tools for these tasks, and for many other core AI tasks. AI has also changed from its focus on deep reasoning to large-scale statistical methods. This partly reflects the huge drop in the cost of memory and processing, along with the availability of networking. So, for example, an area like natural language translation has shifted from an early focus on user modeling and deep reasoning to statistical techniques for machine learning that makes use of large corpus data, particularly text which occurs naturally in online materials such as books, newspapers, social media sites, Wikipedia…. Where early work often involved complex reasoning, now it is possible, and sensible, to explore far simpler methods that harness huge amounts of data to achieve more robust and practical systems.

AI has earned a place as part of a standard computing undergraduate degree. Similarly, some other core areas of the computing syllabus include databases, HCI, software engineering, graphics. Such areas have now established a substantial collection of techniques that belong in the computing professional's toolkit. All of these, not just AI methods, should be used to achieve the core goals of AIED.

AI has achieved much in its long history, often resulting in new communities that are more problem-, rather than technique-focused. For example, robotics researchers have their own publication venues; while they may also publish in AI venues when they create a new contribution to the body of knowledge in AI, their core goals are to create effective robots. High impact research may be based on new ways to make effective use of *existing* software tools for AI, database, graphical, language, vision systems…. Similarly, separate communities have emerged in areas that are central to the AIED vision of effective interfaces, notably natural language generation and understanding and systems based on vision and depth sensing to provide NUI, natural user interaction. This offers support for learning away from the desktop. It opens possibilities for just-in-time learning, teachable moments and kinesthetic interaction that can be valuable for learning.

In summary, AI is pervasive and it is just one, of many, software tools that AIED researchers should draw upon to create the future of technology to enhance learning and education.

3 How has education changed since the birth and naming of AIED?

Over the history of AIED, computing has changed radically. Every potential learner in the developed world now has easy access to many forms of computers in their daily lives. And they will have many more, including personal devices, wearables, mobiles, portables and desktops and well as embedded systems such as interactive tables and walls and smart environments. The interface will have input modalities that include natural language, speech, gaze and gestures as well as keyboard and mouse. Diverse sensors will provide indirect input, such as eye-tracking, mood detectors and activity trackers. Even in the developing world, there is increasing availability of personal technology, particularly mobile phones.

This explosion of computing devices has finally begun to have a deep impact on education, both formal and informal. Our educational institutions make extensive use of computers. Those uses range from core productivity tools, through to tools for particular disciplines as well as personalized and collaborative learning tools. They link the formal and informal, for life-wide learning support.

This has seen the emergence of communities that follow the AIED founder vision for using technology to enhance education. One recent example has seen the emergence of the *Learning Analytics* community. They represent the mainstream of education exploring ways to harness data from even administrative tools (such as those used to capture details of student demographics) and certainly for widespread learning tools, such as Learner Management Systems.

Another emerging example, this time for lifelong, life-wide learning is due to *sensor technology*. For example, wearable activity trackers can be viewed as a valuable data source to an AIED system. They are a form of the interface element, just as surely as a keyboard, drawing tablet or spoken input is. Such sensors can play a key role for personalized teaching, such as interfaces to help people set effective goals and plans, self-monitor progress on these, discover which personal strategies are effective for achieving goals and to learn about new strategies.

Yet another recent EdTech innovation is the *MOOC*. This is exciting on several levels. MOOCs offer the possibility for a very broad population of learners to have access to high quality personalised learning opportunities. MOOC platforms emerged from the elite computer science research world. This is striking as computer scientists, with outstanding expertise in diverse areas of computing, have so clearly committed to creating innovative teaching systems. MOOCs provide exciting green fields for EDM and for translating our years of AIED research into widely used software systems.

These illustrate just three of many trends that matter for AIED. They are pervasive and have high impact. All are currently outside the core of what some members of our community see as AIED. There is a real risk that a paper reporting any of these would be rejected for lacking AI. And authors may assume this, and submit such work elsewhere. Yet all three do offer personalized learning, as the term is described in the broader community. All have data about learners and it is widely recognized that this data is important for informing the learning. Should we call that data a learner model? Why not? Do those communities consider it a learner model? Probably not. Should we object to calling such data a learner model representation just because it is simple by AI standards, rather than complex. Surely these classes of EdTech are within the scope of the vision of the AIED founders.

4 How has AIED changed? And not? Personal case studies.

The last section suggested that AIED has not changed enough to keep up with the dramatic shifts in the real world of education. This section explores some of the ways that the AIED community has already made steps towards accepting research that has little or no AI. There have been AIED papers dealing with essentially the software engineering aspects of sophisticated AI systems. For example, these include the creation of interfaces to make it easier for non-technical users to design and modify the teaching in a complex AIED system; such work tackles the problem that an AIED system needs a better user-friendly interface.

But there has also been work that has no element of AI at all. Lest I risk offending others, I illustrate this in terms of my own work that has been published in AIED and ITS venues but does not have AI. As a young researcher, I was excited at the AIED vision of creating personalized teaching system. I concluded that a key is the learner model because it drives the personalisation, based on its data about the learner. But I was also committed to treating the learner model as the personal data of the learner and to respect the asymmetry in the relationship that should exits between a person and a machine, where the person should be able to maintain a sense of control.

This focus led me to work on creating learner models that respected the learner's right to *control* their own data, to help the learner to be *responsible* for their own learning. As a foundation for learner control, I concluded that it was important to create learner modeling middleware that was designed, from its foundations, to enable the learner to *scrutinize* the *learner model* and the associated *personalization processes*. Issues of personal data privacy are not mainstream AI concerns. But they are important for real world deployments. This is reflected in the 2012 workshop by leaders of the MOOC community, resulting in the Asilomar Convention for Learning Research in Higher Education[2]. While the philosophical standpoint of learner control was a key driver for my research, there are also more pragmatic aspects. One relates to the deeply fallible process of learner modeling. Since the data about learners is generally noisy, unreliable and incomplete, I wanted to create interfaces to the learner model, Open Learner Models (OLMs), that enabled the learner to see their model and how teaching applications interpret and use it. This could enable them to correct it. They could also alter it in other ways if they wished to introduce incorrect data. (The underlying representation avoids this from corrupting the model, and supports multiple views and interpretations of the model). That work was accepted by the AIED and ITS communities, as evidenced by

[2] http://asilomar-highered.info/

publications, such as Kay (2000; 2000a), Kay & Lum (2005) and Czarkowski et al (2005). The learner model representations in that work did not require, or make use of, sophisticated AI.

Concerns for *systems* aspects led to my work on user and learner model servers. This is important for practical systems, but it is not AI (Kay et al 2002; Brusilovsky, 2003; Brusilovsky et al, 2005; Assad et al, 2007; Kay and Kummerfeld, 2012). Designing OLM interfaces is essentially HCI, with a strong focus on user-centred design, rather than AI. The challenge of building systems that work effectively also makes it desirable to create the simplest technical solution that is effective, in that it achieves the intended task. This is good software engineering, good sense and also an excellent foundation for creating OLM interfaces that are simple enough make the model understandable and scrutable. In line with the view of learning data as belonging to the user and under their control, even my earliest implementations of the learner model placed it outside any single application (Kay 1994). The move to learner model servers (Kay et al, 2002) continued the move towards a cloud-based independent learner model as a first class citizen (Kay 2008; Bull and Kay 2010). None of these concerns are AI.

Learner models are clearly core to AIED; they are one of the four elements of personalized teaching. Papers on OLMs have been published in our journal and conferences, as reviewed by Desmarais and Baker (2012). Some have used sophisticated AIED representations, such as cognitive and constraint-based models and Bayesian nets. However, my own work, and key work by other prominent OLM researchers has typically had rather simple learner models. There was no need for complex AI techniques. The defining characteristic of an OLM is that it provides an interface onto a data structure where both were explicitly designed to provide a view of the learner model that would be useful to the learner.

A foundation for designing a learner model is the definition of the domain ontology and the processes to transform learning data into inferences about that learning ontology. In my work, it could more accurately be described as defining the curriculum in terms of the learning objectives. Then the inference is essentially a mapping from learning data onto that curriculum, using the simplest effective interpretation. While some reviewers have criticized some of this work for the lack of AI, they have never explained why a more complex AI approach would be useful or how such modest and simple approaches are inadequate to the task. Nor have they argued the work is not useful. I believe that OLM research is true to the aspirations of the founders of the AIED community, even if it has no element of what is currently AI.

While OLM research is accepted in AIED, my other current research involves creating interfaces for surface computing, with large screen interactive tabletops and walls. This is exciting stuff. Some of it has made it into AIED venues (Martinez-Maldonado et al, 2011, 2012, 2013, 2014). This work used the data from small group interaction at a tabletop to model the effectiveness of collaboration. This used EDM methods to interpret the raw data, to distinguish more, and less, effective collaboration in groups of students. We trialled that work in a lab setting. However, when we moved into the wild, with real classrooms and real teachers, the actual demands of the classroom called for far simpler learner models. For this real world context, we took the same digital footprints of the learners, but this time presented them in very simple OLMs (Martinez-Maldonado et al, 2012, 2014). That was what met the teacher's needs; it did not have or need AI for the core of the research. Some of it seemed to have enough AI or OLM content to make to our conferences, much did not.

In summary, the publications of the AIED community already include some research that provides innovative teaching systems but does not need AI and reports none. But we still exclude other interesting and innovative work, or authors self-exclude it.

5 Summary

This position paper has argued that the foundation vision for AIED was to create personalised learning systems, with highly effective interfaces, and that this vision is still relevant to the AIED community. There is much that remains to be done if we are to create the four core components of AIED architectures. But over the last 25 years, AI has changed, as has education and EdTech. We run the real risk of being left behind some of the most exciting and novel directions if we insist on restricting our research to systems that create or use AI, as it is understood today.

References

1. Assad, M., Carmichael, D. J., Kay, J., & Kummerfeld, B. (2007). PersonisAD: Distributed, active, scrutable model framework for context-aware services. In *Pervasive Computing* (pp. 55-72). Springer Berlin Heidelberg.
2. Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational researcher*, 4-16.
3. Brusilovsky, P. (2003). A component-based distributed architecture for adaptive webbased education. In *Artificial intelligence in education* (pp. 386-388).
4. Brusilovsky, P., Sosnovsky, S., & Shcherbinina, O. (2005). User modeling in a distributed e-learning architecture. In *User Modeling 2005* (pp. 387-391). Springer Berlin Heidelberg.
5. Bull, S., & Kay, J. (2007). Student models that invite the learner in: The SMILI:() Open learner modelling framework. *International Journal of Artificial Intelligence in Education*, *17*(2), 89-120.
6. Bull, S., & Kay, J. (2010). Open learner models. In *Advances in intelligent tutoring systems* (pp. 301-322). Springer Berlin Heidelberg.
7. Czarkowski, M., Kay, J., & Potts, S. (2005). Scrutability as a core interface element. In *Proceedings of the 2005 conference on Artificial Intelligence in Education: Supporting Learning through Intelligent and Socially Informed Technology* (pp. 783-785). IOS Press.
8. Desmarais, M. C., & Baker, R. S. (2012). A review of recent advances in learner and skill modeling in intelligent learning environments. *User Modeling and User-Adapted Interaction*, *22*(1-2), 9-38.
9. Kay, J. (1994). The um toolkit for cooperative user modelling. *User Modeling and User-Adapted Interaction*, *4*(3), 149-196.
10. Kay, J. (2000). Stereotypes, student models and scrutability. In *Intelligent Tutoring Systems* (pp. 19-30). Springer Berlin Heidelberg.
11. Kay, J. (2000a). Accretion representation for scrutable student modelling. In *Intelligent Tutoring Systems* (pp. 514-523). Springer Berlin Heidelberg.
12. Kay, J. (2008). Lifelong learner modeling for lifelong personalized pervasive learning. *Learning Technologies, IEEE Transactions on*, *1*(4), 215-228.
13. Kay, J., Kummerfeld, B., & Lauder, P. (2002). Personis: a server for user models. In *Adaptive Hypermedia and Adaptive Web-Based Systems* (pp. 203-212). Springer Berlin Heidelberg.
14. Kay, J., & Kummerfeld, B. (2012). Creating personalized systems that people can scrutinize and control: Drivers, principles and experience. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, *2*(4), 24.
15. Kay, J., & Lum, A. (2005). Exploiting Readily Available Web Data for Scrutable Student Models. In *AIED* (pp. 338-345).

16. Martinez, R., Wallace, J. R., Kay, J., & Yacef, K. (2011). Modelling and identifying collaborative situations in a collocated multi-display groupware setting. In *Artificial Intelligence in Education* (pp. 196-204). Springer Berlin Heidelberg.
17. Maldonado, R. M., Kay, J., Yacef, K., & Schwendimann, B. (2012). An interactive teacher's dashboard for monitoring groups in a multi-tabletop learning environment. In *Intelligent Tutoring Systems* (pp. 482-492). Springer Berlin Heidelberg.
18. Martinez-Maldonado, R., Kay, J., & Yacef, K. (2013). An automatic approach for mining patterns of collaboration around an interactive tabletop. In *Artificial Intelligence in Education* (pp. 101-110). Springer Berlin Heidelberg.
19. Martinez-Maldonado, R., Clayphan, A., Yacef, K., & Kay, J. (2014). Towards Providing Notifications to Enhance Teacher's Awareness in the Classroom. In *Intelligent Tutoring Systems* (pp. 510-515). Springer International Publishing.
20. Self, J. (2015) The birth of IJAIED, International Journal of Artificial Intelligence in Education. To appear.

A is for Adaptivity, but What is Adaptivity? Re-Defining the Field of AIED

Vincent Aleven

Human-Computer Interaction Institute
Carnegie Mellon University
5000 Forbes Ave
Pittsburgh, PA 15213

aleven@cs.cmu.edu

Abstract This paper proposes to define the field currently known as AIED not in terms of the technology used, but in terms of system behavior. Specifically, it is proposed that AIED is the science and engineering of systems that adapt to learners, so as to help bring about effective, efficient, and enjoyable learning experiences. But what, in general, is adaptivity? Intuitively, being adaptive means that the system adjusts the course of instruction in nuanced and effective ways based on learner differences, for example the goals and needs of individual learners and group of learners. It is difficult to state necessary and sufficient conditions for the concept of adaptivity. Instead, I stipulate that a system is more adaptive to the degree that: (a) its design is grounded in a thorough (empirical) understanding of learners in the given task domain, (b) it is appropriately interactive, and (c) it takes into account, in its pedagogical decision making, how individual learners measure up along different psychological dimensions. These factors help in comparing systems in terms of their degree of adaptivity. They imply that the presence of Artificial Intelligence technology is not a defining factor, even if it can be (and often is) instrumental in bringing about adaptivity.

Introduction

How we define our field (currently called AIED) influences how we position it vis-à-vis other efforts to create learning technologies. This positioning is not merely academic. It may influence public perception and acceptance of our technologies. For example, it may influence how MOOC developers see the need for AIED technology in their courses, and may influence how the technology is accepted and spreads. Experts do not agree about how to define AIED or (relatedly) intelligent tutoring systems [25, p. 21], so the issue is not straightforward. How can we define our field in a way that is inclusive and honors its interdisciplinary nature, while also honoring the range of technologies that are typically being applied, whether AI technologies or not?

As with all educational technology, the goal of our field is to develop a science and practice for the design and implementation of technologies that can support effective, efficient, and pleasurable learning experiences for learners, groups of learners, instructors, and other stakeholders in the educational process. What sets our field apart is that we strive to make our systems "intelligent" or "adaptive," so as to be highly effective with a very wide range of learners. But what do these terms mean? Although the notion of intelligent and adaptive educational technologies is ill-defined, a shared intuition among researchers and practitioners may be that in order to be considered adaptive, a system must be sensitive to important learner differences; a system must have a nuanced way of deciding what, for a given learner or team of learners in a given situation, might be the best way of supporting them, given their learning history and learning goals. Such systems "understand learners" or, more broadly, "care," as John Self famously argued [18].

Artificial Intelligence (AI) can often contribute to creating such systems. It has brought to our field a focus on representation and reasoning, and has highlighted modeling and investigations into the nature of knowledge as a key emphasis in the early days of AIED and intelligent tutoring systems (e.g., [19, 24]). Nonetheless, in my opinion, our field cannot and should not be defined in terms of whether the system has AI or not. One problem is that AI is an ill-defined concept – so it would merely be replacing one ill-defined concept ("adaptive learning technologies") with another. More importantly, AI is neither necessary nor sufficient in order for learning technologies to be adaptive. The use of AI does not in and of itself make a system adaptive in a manner that supports learners effectively. Conversely, not all systems that are adaptive use AI. Also, defining our technologies in terms of the underlying technology seems fundamentally to be barking up the wrong tree. What matters is how learning is supported and whether learning is supported effectively. This viewpoint implies a focus on the behavior of systems [22] much more so than the underlying technology. The question whether AI to stay married to Ed is an interesting one. Perhaps this marriage, which started out so interestingly, needs to now become an open marriage. Better yet, perhaps it needs to be reconceptualized, replaced with a broader, more productive vision, with a renewal of the vows! Definitely, AI should and will remain a central aspect of what we do but it should not be the defining characteristic.

Intuitively, What is Adaptivity?

Proposing that adaptivity should be the defining characteristic of AIED system begs the question, what is adaptivity? Intuitively, we assume that learners differ along (possibly) many dimensions (e.g., prior knowledge, affect, self-regulated learning skills) and that, all else being equal, instruction that takes these differences into account tends to be more effective than instruction that treats all learners as the same. Adaptivity is not binary, something a learning environment either has or does not have. Adaptivity is a matter of degree. Below I offer a more formal definition of adaptivity, first presented in Aleven, Beal, and Graesser [4]. The discussion in the current paper discussion in a paper currently under review [6], although it also broadens and elaborates that discussion. Before I do so, perhaps it helps to get some obvious examples and non-examples on the table. We can then look at more borderline cases and offer a general definition for what it means for a system to be adaptive.

Obvious (i.e., non-controversial) non-examples of adaptive learning technologies are for example textbook problems with final answers to each problem in the back of the book, especially when every student in the same class is assigned the same problems. Other examples that are probably not controversial are online text, lectures with Powerpoint slides, video lectures of famous professors, and documentaries. I am not claiming that these types of instructional material have no place in the educational process [14, 17]. They very well may but they seem to lack adaptivity.

An obvious example of an adaptive learning technology may be an intelligent tutoring system, but what is it that makes it adaptive? A typical answer from our field may be, a rich student model with many student-related variables (knowledge, affect, metacognition, motivation, social factors), updated in real-time, in a sophisticated manner, inferring the unobservable from the observable, and used in sophisticated pedagogical decision making at multiple levels. Each learner or team of learners gets the instruction that is most effective, efficient, or pleasurable for them. Instructional decisions are always based on nuanced, fully up-to-date information.

It may be relevant also to point out that in many discussions about MOOCs and e-learning, a very low bar is used when talking about personalization or adaptivity. For example, Daphne Koller, one of the Coursera co-founders, in her Ted Talk (https://www.youtube.com/watch?v=U6FvJ6jMGHU), hails the ability to provide an error-specific feedback message (on an error discovered through data mining) as an important aspect of personalization of instruction in MOOCs. Further, in a widely-used learning management system such as Moodle (https://moodle.org/) [16], even simple branching structures are considered to be adaptive forms of instruction, in contrast to the intuitions of many ITS researchers.

A Somewhat Unsatisfactory Way to Define Adaptivity?

Let me now examine a prior proposed definition of our concept of interest. The argument has been put forward that a key criterion for adaptivity in learning technologies is that the system has an inner loop [22], meaning that it provides step-level guidance during complex, multi-step problem solving or dialogues. This form of guidance is to be contrasted with answer-level guidance, in which feedback is provided only at the end of each problem. In his 2006 paper, VanLehn views the presence of an inner loop as a defining criterion for intelligent tutoring systems: "Systems that lack an inner loop are generally called Computer-Aided Instruction (CAI), Computer-Based Training (CBT) or Web-Based Homework (WBH). Systems that do have an inner loop are called Intelligent Tutoring Systems (ITS)" [22, p. 233]. In a later article [21], however, he seemed to back off: "Most intelligent tutoring systems have step-based or substep-based granularities of interaction, whereas *most other tutoring systems* [emphasis added] (often called CAI, CBT, or CAL systems) have answer-based user interfaces." Importantly, he points out that systems that provide step-based tutoring tend to have a stronger positive effect on student learning outcomes, compared to no tutoring conditions (i.e., a greater effect size) than systems that provide answer-based tutoring (i.e., do not have an inner loop). VanLehn's definition is attractive in many ways: It emphasizes adaptive behavior as a hallmark of intelligence, which seems right to us. It avoids debates about system architectures or about the thorny question, what is AI? It aligns with key empirical evidence. On the other hand, it is not without its shortcomings, reason perhaps that VanLehn seems to have backed off. Step-based guidance may not be very adaptive if the tutor can only recognize one particular set of steps through each problem. Also, certain desirable forms of adaptivity may not easily be viewed as step-level support (e.g., reacting to student affect or adaptive selection of problems in the system's outer loop). Also, some systems that are commonly considered intelligent or adaptive have rather minimal inner loops such as ASSISTments [12], Wayang Outpost/Mathsprings [9], and Hint Factory tutors [20]. These systems all have a legitimate claim to being adaptive and intelligent. ASSISTments and Wayang Outpost/Mathsprings may not have an elaborate inner loop, but they have other features, such as being designed with a fundamental and sound understanding of student learning. Also, Wayang Outpost in its outer loop adapts to student metacognition and affect in certain ways. Similarly, Hint Factory tutors do not have on-board intelligence, yet behave like an intelligent tutor because of the next-step hint capability.

In this discussion, it is interesting to consider the degree to which specific forms of adaptivity are supported by empirical investigations (e.g., task analysis) and/or rigorous research. For example, step-level feedback and cognitive mastery are strongly supported in the empirical ITS literature, as enhancing student learning [7, 8, 11, 15]. Although the ability to support multiple student strategies within a given problem is widely viewed as desirable, the only study I know that tested this assumption did not find evidence to support it [23].

Adaptivity: A Proposed Definition

Given these considerations, let me now highlight an alternative definition of adaptivity, first presented in a recent article by Aleven, Beal, and Graesser [4], who listed three key elements of advanced learning technologies. For purposes of the current discussion, we can take this term to be synonymous with AIED; the key elements can therefore be viewed of key elements of the kind of adaptivity or intelligence we would like to see in our smart systems for education.

> "Although defining ALTs (advanced learning technologies) is difficult, ALTs have 3 key elements to varying degrees:
>
> - First, these technologies are created by designers who have a substantial theoretical and empirical understanding of learners, learning, and the targeted subject matter.
> - Second, these systems provide a high degree of interactivity, reflecting a view of learning as a complex, constructive

activity on the part of learners that can be enhanced with detailed, adaptive guidance.

- Third, the system is capable of assessing learners, while they use the system, along different psychological dimensions, such as mastery of the targeted domain knowledge, application of learning strategies, and experiences of affective states. On the basis of these assessments, the systems make pedagogical decisions that attempt to adapt to the needs of individual learners."

This definition lists factors, rather than necessary and sufficient conditions, thus acknowledging that adaptivity is an open-textured concept, that is, a concept whose meaning needs to be interpreted as we go, perhaps on a case-by-case basis, and perhaps with a shift in meaning over time, as our field evolves and develops new and innovative forms of instructional support. Listing factors helps with defining the concept flexibly in a way that enables us to talk about degrees of adaptivity, rather than view it as binary. It is interesting to point out, further, that these elements are technology-agnostic; no specific technologies are mentioned or assumed. It is reasonable to think that the second and third key elements (interactivity with detailed guidance based on learner variables assessed by the system) will often involve AI technology. AI might be a particularly good match, given its emphasis on knowledge representation, reasoning, and problem solving, its concerns with diagnostic processes needed to infer and update learner models, and its concern with the nature of knowledge to be learned (e.g., [24]). Nonetheless, AI cannot be the one defining ingredient of what makes our systems adaptive.

On a personal note, this definition marks an expected return to a central theme of my dissertation, which dealt with a tutoring system, CATO, for case-based legal argumentation, a quintessential ill-defined task domain [1, 2, 3]. CATO was designed to help beginning law students learn skills of argument by analogy, a common form of argument in the legal domain. That is, this work addresses debates about whether a given new case (a problem situation about which a legal claim has arisen) properly belongs to an open-textured category, which, as in our current discussion, was defined by factors, rather than necessary and sufficient conditions. A key mode of analyzing, exploring, and arguing is to compare the new case to carefully selected past cases with favorable and unfavorable decisions [10], with the factors functioning as key dimensions of comparison. In the legal domain, comparisons with past cases that have been authoritatively classified often bring substantial clarity, although not often provably correct answers. And so it is with our question of what it means for a learning environment to be adaptive, although with an interesting twist: Our own domain lacks authoritative classifications; we do not have a supreme arbiter of whether systems are officially AIED systems or not (nor, of course, should we strive to have such an arbiter). We do have paradigm cases, however, landmark intelligent tutoring systems and even the hypothetical intelligent tutoring system sketched above. These systems can play an important role as anchors in enlightened discussions about the foundations of our field.

Element 1: Design Based on an Empirically-Grounded Understanding of Learners

Perhaps it helps to elaborate on each of the three key elements (or factors). Interestingly, the first element (i.e., the requirement that the designers have "a substantial theoretical and empirical understanding of learners, learning, and the targeted subject matter") relates to the *design* of the system, not to system features or techniques/methods/algorithms under the hood. (The discussion of this factor is informed by debates I have had with my colleague Ken Koedinger.) This requirement could be met in many different ways. Specific to the concerns of the field of AIED, the first part of this definition emphasizes (implicitly) the use of cognitive task analysis and educational data mining to guide system design or redesign, development, and cyclical improvement. A particularly attractive scenario is that the designers carry out cognitive task analysis activities up front to study learners' ways of thinking in the given domain including their strategies and informal shortcuts, but also including the specific conceptual and procedural difficulties they experience. This scenario continues with the data-driven refinement of the system, preferably in ways in which the overall effectiveness of the system, in terms of out-of-system transfer of learning outcomes, preparation for future learning, learner (and instructor) satisfaction, and so forth, is continuously assessed, so that improvement from cycle to cycle is clearly visible. It may be clear that this vision fits particularly well with the current emphasis of big data in education. The fields of EDM and AIED can be at the forefront of this movement (see, e.g., [5]).

A somewhat different way of thinking about this requirement may be that the project team has specialists in a variety of fields, not just technology experts but also researchers in relevant branches of psychology, in education in the given subject area (e.g., math education, science education, legal reasoning, and so forth), as well practitioners.

This first factor implies a substantial broadening of how we think about adaptivity, compared for example to the intuitive notion discussed above and more generally, compared to how we, as a field, have construed the notion of adaptivity up until now. It raises the possibility of considering the design of systems, including even the choice of problems sets and detailed learning objectives, as part of what makes a system adaptive. It may even make it possible to see a modicum of adaptivity in some of our prime examples of instructional materials previously considered as non-controversial non-examples, such video lectures. When designed to target known challenges in learning, they meet the first factor, the more so when based on extensive empirical investigations of what is hard for learners to learn. They would however not be strong examples, as they would not meet the second and third factors.

Element 2: Interactivity

The second requirement for adaptivity is that a system supports a high degree of interactivity, to provide guidance in complex and constructive learning activities. I do not mean to say that more interactivity is always better; rather, in emphasizing the adaptive nature of the guidance that the system gives, the system is capable of providing an appropriate amount of guidance for the given learner(s) at the specific junction in their learning process. How much guidance is appropriate at what stages of learning is an interesting question [13].

The second factor was included to help capture the emphasis that our field places on constructive learning activities and on learning by doing, rather than learning by (merely) reading, watching or listening. An interesting data mining study of data from a psychology online course suggests that learning by doing yields six times greater learning than reading online text in the course or watching the video lectures [14]. A clear cut case of the second factor would be an intelligent tutoring systems with detailed guidance in their inner loop, even if we do not consider the presence of an inner loop as a defining characteristic. I do not mean to rule out systems or projects that focus on enhancing reading, watching, or listening by means of interactive support for comprehension or metacognitive strategies, for example. The second factor was included partly to help rule out (or at least, help view as low on the adaptivity scale) the non-controversial non-examples listed above, such as fixed problem sets with only answer-level feedback in the back of the book, or long video lectures without interactive activities

Frankly, this second factor is the factor that I am the least sanguine about; it may be somewhat redundant with the third factor, and it is difficult to view interactivity per se as a good thing, contributing to learning. Then again, discussions around the notion of interactivity are interesting, as long as the discussants are mindful that it is not interactivity per se that matters, but how it supports learning or other desirable educational outcomes. Further, this factor highlights an important connection, namely, that of our field with the broader field of human-computer interaction.

Element 3 –Change Instruction Depending on Learner Differences

The third requirement, as mentioned, is that the system in its pedagogical decisions takes into account that learners differ and that the same learner is not the same for very long; learners change as they learn. For example, different learners have different prior knowledge, may experience different affect during a given learning task, tend to have different goals for learning the material, and may differ in how they regulate their own learning. In collaborative learning situations, learners may have different collaboration skills and social skills; they may be a good or a poor match regarding prior knowledge or personality, and so forth. A system should be considered as more adaptive to the extent that it adjusts its instruction, both in the inner and outer loop, based on these learner variables and perhaps others.

This requirement is consistent with Woolf's emphasis on a system having a student model and using it to adapt instruction [25], traditionally viewed as a hallmark of "intelligent" tutoring. The system builds up and maintains a student model by continuously assessing learners along various psychological dimensions (cognitive, metacognitive, motivational, and so forth). This student model is then used as the basis for individualization. Perhaps the requirement that it is the system doing the assessing is too stringent. Perhaps the viewpoint that what is being assessed is the learner is too stringent as well. Alternative viewpoints would be that a group of learners is being assessed or perhaps that the system interprets the situation more than the learner(s) or group of learners, if that distinction makes sense (it may not). I do not mean to argue, however, that we define our field in terms of whether or not systems have a student model. That is, I do not mean to equate AIED with the field of UMAP. For example, it is conceivable that a system could be strong with respect to the first two factors but not the third and be generally accepted as belonging to the field of AIED.

There are many interesting open questions regarding how systems (as well learning environments not strongly supported by technologies) *should* adapt to learners and which learner variables (or learner group variables) are most important in this regard. In my opinion, our field is uniquely positioned to extend the science of how instruction should adapt to individual differences. Of the three factors, the third reflects most clearly how we have traditionally viewed our field.

Final Remarks

In closing, it may be worth re-iterating that the proposed definition of adaptivity does not place emphasis on particular technologies; rather, it emphasizes the behavior of systems, much in line with VanLehn's seminal 2006 article [22] and also in line with the Turing test as a behavioral test of intelligence. Another attractive property of this definition is that also honors the interdisciplinary foundations of our field. In my view, AIED was never only about technology (CS/AI, computational linguistics, and so forth); its strength has always been that it included people and methodologies from different fields, such as human-computer interaction, psychology (cognitive, educational, developmental, social), education, design, statistics, and so forth. The field and its methodologies are interdisciplinary. Empirical evaluation of systems building has always been highly valued in our field, sometimes even to a fault (e.g., when interesting new technology developments were not given air time at conferences before there are proven results). The emphasis on high-quality empirical work is enormously important toward the goal of creating a science for the design and implementation of technologies that can support effective, efficient, and pleasurable learning experiences for a wide range of learners.

An implication of the proposed definition is that reviewer comments that "there is no AI in the system" or "the work does not push the envelop in terms of AI algorithms applied to education" should be a thing of the past. Instead, reviewer feedback should refer to the factors listed above: systems not being designed with deep insight into learning and learners' difficulties, not being interactive, and not being able to react in nuanced ways that make learning better.

The way for AI to stay married to Ed is perhaps not to declare it an open marriage, but rather, to re-define the marriage so it is appropriately broad and open-ended, a way of renewing the vows. We hope that the thoughts offered in this paper can be helpful.

Finally, what's in a name? A lot, I would argue. Our name reflects how we view ourselves, and in turn, how the rest of the world views us. Our current name honors AI as a central component as we do. I would much prefer that the disciplinary diversity and focus on behavior of systems be central. How about:

AIED = Adaptive Instruction: Evaluation and Design?

Or, if we are willing to tolerate AIEDD, how about:

AIEDD = Adaptive Instruction: Evaluation, Development, and Design

Acknowledgments

This paper benefited from discussions with my colleague Ken Koedinger as well as from the helpful encouragement and pushback from the anonymous reviewers.

References

1. Aleven V.: Using background knowledge in case-based legal reasoning: A computational model and an intelligent learning environment. Artificial Intelligence 150, 183-237 (2003)
2. Aleven V.: An intelligent learning environment for case-based argumentation. Technology, Instruction, Cognition, and Learning 4, 191-241 (2006)
3. Aleven V., Ashley K.D.: Teaching case-based argumentation through a model and examples empirical evaluation of an intelligent learning environment. In: du Boulay B., Mizoguchi R. (eds.) Artificial Intelligence in Education, Proceedings of AI-ED 97, pp. 87-94. IOS Press, Amsterdam (1997)
4. Aleven V., Beal C.R., Graesser A.C.: Introduction to the special issue on advanced learning technologies. Journal of Educational Psychology 105, 929-931 (2013, Nov)
5. Aleven V., Koedinger K.R.: Knowledge component approaches to learner modeling. In: Sottilare R., Graesser A., Hu X., Holden H. (eds.) Design recommendations for adaptive intelligent tutoring systems, pp. 165-182. US Army Research Laboratory, Orlando, FL (2013)
6. Aleven V., McLaren B.M., Sewall J., Popescu O., et al: Towards tutoring at scale: Reflections on "A new paradigm for intelligent tutoring systems: Example-Tracing tutors". International Journal of Artificial Intelligence in Education (under review)
7. Anderson J.R., Conrad F.G., Corbett A.T.: Skill acquisition and the LISP tutor. Cognitive Science 13, 467 - 505 (1989)
8. Arroyo I., Beck J., Woolf B.P., Beal C.R., Schultz K.: Macroadapting AnimalWatch to gender and cognitive differnces with respect to hint interactivity and symbolism. In: Gauthier G., Frasson C., VanLehn K. (eds.) Proceedings of the 5th International Conference on Intelligent Tutoring Systems (ITS 2000), pp. 574-583. Springer Verlag, Berlin (2000)
9. Arroyo I., Woolf B.P., Burleson W., Muldner K., et al: A multimedia adaptive tutoring system for mathematics that addresses cognition, metacognition and affect. International Journal of Artificial Intelligence in Education 24, 387-426 (2014, Dec)

10. Ashley KD. Modeling legal argument: Reasoning with cases and hypotheticals. Cambridge, MA: MIT press; 1991.
11. Corbett A., McLaughlin M., Scarpinatto K.C.: Modeling student knowledge: Cognitive tutors in high school and college. User Modeling and User-Adapted Interaction 10, 81-108 (2000)
12. Heffernan N.T., Heffernan C.L.: 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, 470-497 (2014, Dec)
13. Koedinger, K. R., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. Educational Psychology Review, 19(3), 239-264.
14. Koedinger K.R., Kim J., Jia J.Z., McLaughlin E.A., Bier N.L.: Learning is not a spectator sport: Doing is better than watching for learning from a MOOC. In: Kiczales G., Russell D.M., Woolf B. (eds.) Proceedings of the Second (2015) ACM Conference on Learning @ Scale, pp. 111-120. ACM, New York (2015)
15. McKendree J.: Effective feedback content for tutoring complex skills. Human-Computer. Interaction 5, 381-413 (1990, Dec)
16. Rice W. Moodle 2.0 e-learning course development a complete guide to successful learning using moodle 2011:
17. Schwartz D.L., Bransford J.D.: A time for telling. Cognition and Instruction 16, 475-5223 (1998)
18. Self J.: The defining characteristics of intelligent tutoring systems research: ITSs care, precisely. International Journal of Artificial Intelligence in Education 10, 350-364 (1998)
19. Sleeman D, Brown JS. London: Academic Press; 1982.
20. Stamper J., Eagle M., Barnes T., Croy M.: Experimental evaluation of automatic hint generation for a logic tutor. International Journal of Artificial Intelligence in Education 22, 3-17 (2013)
21. VanLehn K.: The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. Educational Psychologist 46, 197-221 (2011)
22. VanLehn K.: The behavior of tutoring systems. International Journal of Artificial Intelligence in Education 16, 227-265 (2006)
23. Waalkens M., Aleven V., Taatgen N.: Does supporting multiple student strategies lead to greater learning and motivation? Investigating a source of complexity in the architecture of intelligent tutoring systems. Computers & Education 60, 159 - 171 (2013)
24. Wenger E. Artificial intelligence and tutoring systems: Computational and cognitive approaches to the communication of knowledge. Los Altos, CA: Morgan Kaufmann; 1987.
25. 24. Woolf BP. Building intelligent interactive tutors: Student-centered strategies for revolutionizing e-learning. Burlington, MA: Morgan Kaufmann; 2009.

All that glitters (in the lab) may not be gold (in the field)

Amruth N. Kumar

Ramapo College of New Jersey, Mahwah, USA
amruth@ramapo.edu

Abstract. AI-ED community has hewed to rigorous evaluation of software tutors and their features. Most of these evaluations were done *in-ovo* or *in-vivo*. Can the results of these evaluations be replicated in *in-natura* evaluations? In our experience, the evidence for such replication has been mixed. We propose that the features of tutors that are found to be effective *in-ovo*/*in-vivo* might need motivational supports to also be effective *in-natura*. We speculate that some features may not transfer to *in-natura* use even with supports. Recognition of these issues might bridge the gap between AI-ED community and educational community at large.

Keywords: In-ovo, in-vivo, in-natura, replication of results.

1 Introduction

Evaluation of software tutors may be carried out in one of three settings:

- *In-ovo*: Research subjects hand-picked for the evaluation use the software tutor in a laboratory setting, typically under tightly controlled conditions, and under the supervision of the researcher.
- *In-vivo*: Students enrolled in a course use the software tutor in the class room, typically under tightly controlled conditions and under the supervision of the researcher or course instructor.
- *In-natura*: Students enrolled in a course use the software tutors, typically after class, on their own time, and unsupervised.

These three types of evaluation are summarized in Table 1.

Type	Location	Subjects	Conditions	Supervised
In-ovo	Laboratory	Recruited	Controlled	Yes
In-vivo	Classroom	Students enrolled in a course	Controlled	Yes
In-natura	After-class		Not controlled	No

Table 1: Types of evaluation of software tutors

AI-ED community has reported frequently using *in-ovo* and *in-vivo* evaluations in its studies of the effectiveness of software tutors and their features. Researchers have strictly controlled the conditions of these studies – what a subject can do or not do during the study, whether the subject is exposed to any distractions during the study, etc. – so as to minimize the influence of extraneous factors.

However, in real-life, especially at baccalaureate level, software tutors are less used as in-class exercises than as after-class assignments or study aides. The reasons for such use are many, including: course instructors may not want to spend valuable class time using software tutors; and students may not have access to (sufficient numbers of) computers during class.

When software tutors are used for after-class assignments, mandatory or otherwise, issues of intrinsic and extrinsic motivation play a much larger role in their use and utility. For starters, the popular aphorism *If you build it, they will come* does not apply to software tutors – unless students are required to use a software tutor, they will not use it (in any significant numbers). This significantly drives down participation and may skew evaluation results because of the self-selected nature of subjects. When they do use it, extrinsic motivation often plays a larger role than intrinsic motivation – if they are awarded course grade proportional to how well they do on the software tutor, they are more likely to engage seriously with the tutor. On the other hand, if they are given credit simply for using the software tutor, they are likely to do the least amount of work possible to qualify for such credit.

Given these considerations, do the research results elicited under carefully controlled conditions *in-ovo* or *in*-vivo extend to *in-natura* use of software tutors? In other words, can results obtained *in-ovo* or *in-vivo* be replicated *in-natura*? Our experience has been mixed. We will present results from evaluations of two features – reflection and self-explanation - vouched for by the AI-ED community that did not pan out in our *in-natura* evaluations.

For our evaluations, we used software tutors for programming concepts, called problets (problets.org). These tutors are being used every semester by 50-60 schools, both undergraduate and high-school. Since problets are deployed over the web, students have access to the software tutors anytime, anywhere. Problets are set up to automatically administer pre-test-practice-post-test protocol every time they are used [5]. They have been continually used and evaluated *in-natura* since fall 2004.

2 Reflection

The benefits of post-practice reflection have been studied by several researchers (e.g., [3]). In problets, we introduced reflection in the form of a multiple-choice question presented after each problem. The question states "This problem illustrates a concept that I picked based on your learning needs. Identify the concept." The learner is provided five choices, each of which is a different concept in the domain. The learner must select the most appropriate concept on which the problem might be based, and cannot go on to the next problem until (s)/he correctly selects it. The problet records the number of unique concepts selected by the learner up to and including the most appropriate concept. See Figure 1 for a snapshot of the reflection question presented after the student has solved a problem on selection statements.

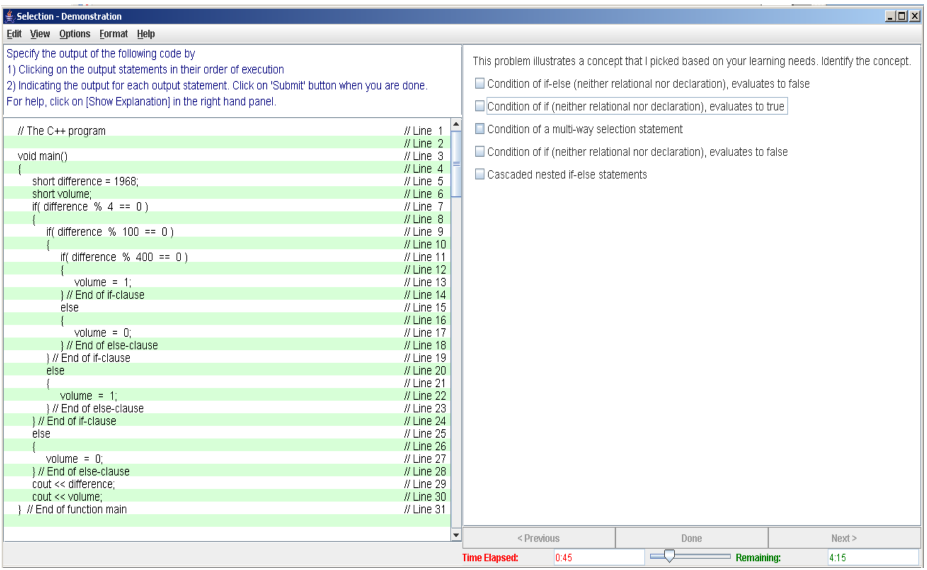


Figure 1: Selection tutor: Problem in the left panel; Reflection question in the right panel

We conducted several controlled evaluations of reflection [8] using selection and while loop tutors in 2006-07. Control group was never presented any reflection questions. Test group was presented a reflection question after each problem during pre-test, practice and post-test. If a student solved a problem incorrectly, the student was required to answer the subsequent reflection question correctly before going on to the next problem.

Practice was adaptive, and based on the student's performance on the pre-test. The entire protocol was limited to 30 minutes for control group and 33 minutes for test group. For analysis purposes, we considered only *practiced* concepts [5], i.e., concepts on which the student solved a problem incorrectly during pre-test, solved one or more problems during adaptive practice and also solved the post-test problem before running out of time.

Table 1 lists the score per problem on pre-test and post-test of all *practiced* concepts. No significant difference was found between control and test groups, indicating that the two groups were comparable. However, no significant difference was found in their pre-post improvement either, suggesting no differential effect of reflection on their learning. Please see [8] for additional details of the evaluation.

Score per problem	Pre-Test	Post-Test	Pre-post p
Control Group (Without Reflection) (N =89)			
Mean	0.118	0.736	< 0.001
Standard-Deviation	0.177	0.353	
Test Group (With Reflection) (N =152)			
Mean	0.144	0.787	< 0.001
Standard-Deviation	0.183	0.319	
Between groups p	0.283	0.266	

Table 1: Both the groups improved significantly from pre-test to post-test; the difference between the two groups was not significant on either the pre-test or the post-test

3 Self-Explanation

The effectiveness of providing self-explanation questions in worked examples has been well documented by AI-ED community (e.g., [1]).

Selection tutor was used for this study. When the student solves a problem incorrectly, the tutor presents feedback including step-by-step explanation of the correct execution of the program in the fashion of a fully worked-out example. Self-explanation questions were presented embedded in this step-by-step explanation, as shown in Figure 2. Each self-explanation question is a drop-down menu that deals with the semantics of the program, e.g., the value of a variable, the line to which control is transferred during execution, etc. The questions were independent of each other, but answering them required the student to closely read the step-by-step explanation/worked out example and understand the behavior of the program in question.

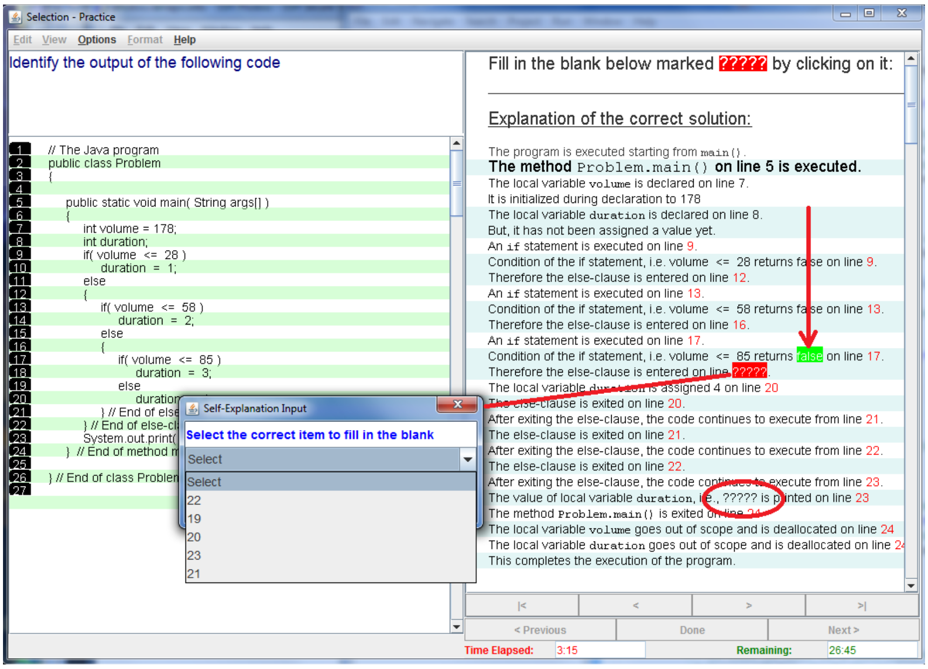


Figure 2. Snapshot of selection tutor with self-explanation questions displayed in the right panel

So as not to overwhelm the student, the tutor limited the number of self-explanation questions per problem to three. The student was allowed as many attempts as needed, but had to answer each self-explanation question correctly before proceeding to the next question, and had to answer all the self-explanation questions correctly before proceeding to the next problem. A version of the tutor was used for the control group that did not present any self-explanation questions. This version of the tutor allowed the learner to advance to the next problem as soon as it displayed step-by-step explanation of the current problem.

Controlled evaluation of selection tutor was conducted *in-natura* over three semesters: fall 2012-fall 2013 [4]. No significant difference was found in the average score per pre-test problem between control (N = 395) and test (N = 335) groups [$F(1,729) = 1.018$, $p = 0.313$]. So, the two groups were equivalent. The mean number of concepts practiced by control group was 1.62, and by test group was 1.78. However, since control group was allowed 30 minutes to practice with the tutor and test group was allowed 40 minutes, univariate analysis of the number of concepts practiced was conducted with self-explanation as the fixed factor and total time spent as the covariate. The difference between the two groups was found to be significant [$F(2,597) = 62.207$, $p < 0.001$]: accounting for the extra time allowed, control group practiced 1.72 ± 0.11 concepts whereas test group practiced 1.662 ± 0.12 concepts. Therefore, test group practiced significantly fewer concepts than control group. No significant difference was found between the two groups on the pre-post change in score on practiced concepts, suggesting no differential effect of self-explanation on learning. Please see [4] for additional details of the evaluation.

4 Discussion

In both the studies – on reflection and self-explanation – we have verified that our implementation is behaviorally similar to, if not the same as described in at least some of the literature on the topic published in the AI-ED community. Even if our interpretation of both reflection and self-explanation behaviors differs enough from those reported in literature to render our treatments ineffective, we would expect that the increased time-on-task due to these faux treatments would have still yielded some learning benefits.

Our evaluations cannot be faulted for inadequate participation – our evaluations have typically involved 200-300 students, which is an order of magnitude larger than the number of subjects reported in typical *in-ovo* and *in-vivo* evaluations.

We have used standard protocols for evaluation – controlled studies, pre-test-practice-post-test protocol and partial crossover design. We have used ANOVA for data analysis. In our studies, we have considered only *practiced* concepts – concepts on which students solved problems during all three stages of the protocol: pre-test, practice and post-test, so noise is not an issue in the analyzed data.

These practices have been effective - not all our evaluations have come up empty, e.g., we have found significant effect of providing error-flagging feedback on test performance (e.g., [6]), and significant stereotype threat (e.g., [7]).

An explanation for the lack of results might be the difference in student motivation in *in-ovo*/*in-vivo* versus *in-natura* evaluation. Apart from issues of extrinsic motivation mentioned earlier, it may also be argued that given the lack of supervision in *in-*

natura evaluation, students are less likely to experience *Hawthorne effect* [2]. So, the features of tutors that are found to be effective in *in-ovo*/*in-vivo* evaluations might need motivational support to also be effective in *in-natura* evaluations.

Then again, even with motivational support, students may resent having to perform tasks (such as answering questions on reflection) that they do not perceive as directly contributing to their assignment at hand, and may not participate in, or may not be amenable to benefiting from what they view as a chore. In other words, some features may not be transferable from the laboratory to the field regardless of the supports provided.

While we have focused on the transferability of evaluation results from lab/classroom to after-class setting, researchers have reported similar issues transferring results from the lab to the classroom, e.g., in a study of politeness in intelligent tutors [9], researchers reported finding weaker results when the study was conducted in a classroom rather than a laboratory. They speculated that grades, an extrinsic motivational factor, may be to blame. Furthermore, they wrote [9], "In the rough-and-tumble of the classroom, with its noise, question-asking, and social environment, students *may simply not concentrate as much on the feedback provided by the computer tutor*. The lab setting, on the other hand, is a quiet environment where subjects work on their own with few distractions, and certainly none from classmates and a teacher" (italics not in the original). The noise, distractions and lack of structure used to describe a classroom as compared to laboratory setting are the very same terms, magnified, that could be used to describe an after-class setting as compared to a classroom. In other words, when it comes to noise, distractions and lack of structure, laboratory and after-class setting are at opposite ends of a spectrum, with the classroom situated in between. That *students may not concentrate as much on the feedback provided by the tutor* may explain why reflection and self-explanation, both provided as part of feedback, failed to live up to expectation in our *in-natura* evaluations.

It appears that *in-natura* use of software tutors entails more than just large-scale/unsupervised deployment of *in-vivo* results and *in-vivo* use entails more than just live-classroom deployment of *in-ovo* results. Motivational supports may be needed to transition results from the laboratory to the field and some results found in the laboratory may fail to transfer to the field even with motivational supports. Treating *in-natura* use of software tutors as being distinct from *in-ovo/in-vivo* uses is reminiscent of the outgrowth of Chemical Engineering as a discipline of the field from Chemistry as a discipline of the laboratory. While Chemistry is the study of properties of materials, Chemical Engineering is the study of the production of materials on an industrial scale, albeit with its basics firmly rooted in Chemistry. In the early years, chemists refused to accept Chemical Engineering as anything more than Chemistry, and engineers refused to recognize Chemical Engineering as an engineering discipline [10], but not so any more. May be AI-ED community should treat *in-natura, in-vivo* and *in-ovo* as three independent, necessary and valuable stages in the evaluation of any treatment. May be, *in-natura* evaluation is what is needed for educational community at large (especially higher-education community) to recognize and incorporate the important pedagogical insights being offered by AI-ED community.

Acknowledgments. Partial support for this work was provided by the National Science Foundation under grants DUE-0817187 and DUE-1432190**.**

5 References

1. Conati, C. and VanLehn, K. (1999) Teaching meta-cognitive skills: implementation and evaluation of a tutoring system to guide self-explanation while learning from examples. Proc. AI-ED 99, 297-304.
2. Franke, R.H. and Kaul, J.D. The Hawthorne experiments: First statistical interpretation. *American Sociological Review.* Vol 43. 1978. 623-643.
3. Katz, S., O'Donnell, G., Kay, H. (2000) An Approach to Analyzing the Role and Structure of Reflective Dialogue. International Journal of Artificial Intelligence in Education, 11, 320-343.
4. Kumar, A.N. An Evaluation of Self-Explanation in a Programming Tutor. In Proc. Of ITS 2014, Hawaii, June 2014. 248-253.
5. Kumar, A.N. A Model for Deploying Software Tutors. IEEE 6th International Conference on Technology for Education (T4E). Amritapuri, India, 12/18-21/2014, 3-9.
6. Kumar, A.N. Limiting the Number of Revisions While Providing Error-Flagging Support During Tests. Proc. Intelligent Tutoring Systems (ITS 2012), LNCS 7315, Chania, Crete, 6/14-18/2012, 524-530.
7. Kumar, A.N. A Study of Stereotype Threat in Computer Science. Proceedings of Innovation and Technology in Computer Science Education (ITiCSE 2012). Haifa, Israel, 7/3-5/2012, 273-278.
8. Kumar, A.N. Promoting Reflection and its Effect on Learning in a Programming Tutor. Proceedings of 22nd International FLAIRS conference on Artificial Intelligence (FLAIRS 2009) Special Track on Intelligent Tutoring Systems, Sanibel Island, FL, May 19-21, 2009, 454-459.
9. McLaren, B.M., DeLeeuw, K.E., and Mayer, R.E. Polite web-based intelligent tutors: Can they improve learning in classrooms? *Computers and Education.* 56(3): 574-584, 2011.
10. Reynolds, Terry S. Engineering, Chemical, in Rothenberg, Marc, *History of Science in United States: An Encyclopedia*, New York City: Garland Publishing, 2001. ISBN 0-8153-0762-4

Why AIED Needs Marriage Counselling by Cognitive Science (to Live Happily Ever After)

Björn Sjödén

Lund University Cognitive Science, Sweden
Bjorn.Sjoden@lucs.lu.se

Abstract. In this position paper, I reflect upon the question "Should AI stay married to Ed?", specifically referring to how research in AI and Education should cross-fertilize to define AIED as an independent practice, beyond its composite fields. In my view, a mix of approaches, inspired by cognitive science, should serve to formulate characteristic research questions for the AIED community. Such questions may be derived from considering the social context of learning and how it is applied in artificial systems, as exemplified by educational games and ITS with Teachable Agents. I conclude by suggesting two discussion points of emergent interest to AIED research: (1) How can we formulate scientifically based guidelines for the use and evaluation of educational software? (2) Is there anything such as "unique AIED competence" and, if so, what does this imply for the AIED identity?

Keywords: AIED, marriage, multidisciplinary research, educational games, ITS, Teachable Agents.

1 Introduction

What *is* and what *should be* the role of AI in Education and conversely of Education in AI?

For all its successes, the very need to reassert AIED's position as a research field after 25 years may reflect two critical shortcomings: a failure to appreciate its relative independence from both AI and education, on the one hand, and an underused and conservative application of AI for educational purposes that has not been fully embraced by educators, on the other. One might compare to fields like HCI or interaction design, which have successfully defined and built research communities around cross-disciplinary domains that focus on people's use of technology.

A stumbling-block to AIED practitioners might be that the field has no obvious "core", that is, it has no clearly defined subject of investigation, such as "computers" or "interactive systems" or even an abstract topic like "instructional strategies". The most concise, official description of the field appears in operational terms, with respect to the scope of the AIED journal, as "the application of artificial intelligence techniques and concepts to the design of systems that support learning" (from ijaied.org). This leaves room for a great variety of research and different approaches – which is good – but the field seems to lack a common conceptual framework for relating advances in AI to advances in education research that would inform characteristic AIED research questions. Is it at all clear for the field's different practitioners what the common denominator of AIED research is?

I posit that it means something to be knowledgeable in AIED and being skilled in AIED research as such, beyond having expertise in AI and Education as distinct fields. The identity of the AIED field is then formed by the content of this "AIED competence", its unique contributions and necessary limitations to other areas. In effect, other considerations become important for AIED than in traditional AI research that does not necessarily apply to education. For example, the AIED researcher with an AI background might be more concerned with "weak AI" as a means to make students learn better or pay more effort, while having to take into account what can be realistically implemented and evaluated in a school or classroom setting, on different technical platforms (tablets, smart phones, laptops etc.) and for different groups of students. Likewise, an AIED researcher with an education or pedagogy background would look to how the use of technology can add to present pedagogical strategies and teaching methods as a means to achieve the same goals. Eventually, as research from both ends cross-fertilize, they may transform educational practices by setting new learning goals defined by the use of technology (e.g. "21st century skills") [1].

In this text, I present a view of AIED that develops from practical considerations for a functioning relationship between AI and Ed, but also forms a new area of research for educational purposes. The educational context both constrains and opens up a largely unexplored scene for novel applications of AI techniques that further motivates the growth of AIED as an independent field. As such, I argue that AIED should aspire to achieve two overarching goals: (1) Improve human learning, and (2) Inform and expand the scientific basis of education. (Notably, the AIED Society has set as its aims to promote knowledge and research in AIED but does not explicate the aims of the field itself.)

As a field of empirical, scientific inquiry (and not just the pragmatic "application of AI techniques"), AIED may be fruitfully compared to Cognitive science. The success of cognitive science as an academic discipline shows how intrinsically different fields – among those psychology, biology, computer science, anthropology and philosophy – have found a common identity in the pursuit of certain well-recognized research questions under the multidisciplinary banner of "cognition". Notably, one does not have to be an expert in all these fields to become an expert cognitive scientist, and it is possible to work within any of these fields without doing research of intrinsic interest to cognitive science. Thus, cognitive science found an identity of its own from combining perspectives and methods from various disciplines, in principle not different from how AIED can develop from merging aspects of AI and education research.

The first question then becomes how the multidisciplinary AIED field should be conceptualized in relation to its history and previous accounts. Second, we need to know what the content of the practice is – the research outcomes and applications – that motivates AI and Ed's relationship. By setting the example, cognitive science might be just the marriage counsellor that AI and Ed need to develop their common interests and secure the future well-being of AIED.

2 Reconceptualizing AIED as a multidisciplinary field

Looking back, Cumming and McDougall [2] already in 2000 speculated how AIED might be "mainstreaming into education" in the (then) future of 2010. They argued both for relabeling the field (“especially the ‘AI’”, p. 204, which they considered does not communicate the field well) and its crucial need for “AI expertise of the highest order” in order to “keep at the forefront of all of the contributing disciplines” (p. 205). This appears, to express it mildly, as a tall order for AIED to take. Above all, it seems to indicate that the field has long had an unclear identity, particularly when it comes to defining the kind of AI expertise needed for being an AIED, rather than an AI or educational, researcher.

I will not propose a new label for AIED research, but perhaps its identity should not be formed on basis of its historical, composite fields, but rather from what motivates AIED research as a multidisciplinary practice in the present and for the future. As to the topics of research, there is a vast array of educational technologies available today that did not exist at the field’s inception 25 years ago. In short, things have changed, and besides emerging new technologies, there is an emerging new generation of AIED researchers.

Looking forward, I approach this question from the perspective of a beginning researcher in the field, who needs to define his future area of expertise. While being actively involved in the AIED community [e.g. 3, 4, 5, 6], I do not see myself as belonging either to the “AI field” or the “Education field” or at least not exclusively so. Rather, and for reasons outlined in the introduction, I would attest to Cumming and McDougall’s [2] observation that “Many AIED researchers would be happy to be described as cognitive scientists.” (p. 198) and their suggestion that AIED “should overlap with cognitive science” (p. 205).

Like cognitive science, albeit on a smaller scale, AIED may play a crucial role for bringing computer science-oriented (AI) and psychology/pedagogy-oriented (Education) research together. Figure 1 illustrates how this view of an emerging AIED field differs from previous conceptions of bringing AI and Ed together.

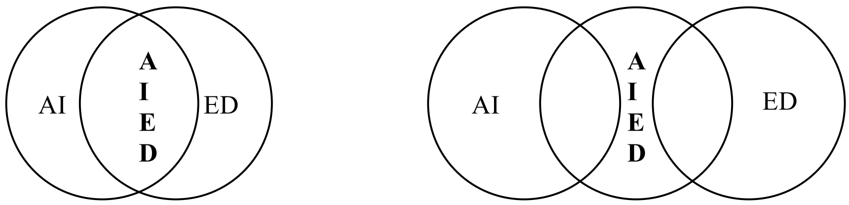


Fig. 1. Two alternative conceptions of AI + Ed: (left) AIED as the combinatory interests of AI and Education research; (right) AIED as an independent, multidisciplinary field, defining its own aims and scope in between the respective fields of AI and Education.

The emancipation of AIED research from its surrounding disciplinary boundaries (Fig. 1, left) would make room for its own defining research questions (Fig. 1, right). Whereas AI puts machine learning and human-like intelligence in focus, Education focuses on fostering human learning and intelligence. AIED knowledge should then serve to bridge this gap by informing techniques to promote more efficient and intelligent interactions with humans that improve educational outcomes (rather than, say, aiming to reproduce human abilities or solving computational problems that do not feed back to students). A combination of methodological approaches is likely needed, as every method carries with it implicit assumptions about what knowledge it can produce (from an AI as well as an education perspective).

Often multidisciplinary knowledge is needed to appreciate the educational applications of different technologies. A classic example would be how computers in school are most often used as word processing and communication tools, whereas the sophistication of the underlying technology would make the computer the obvious arena for students to learn from and by AI technologies in any school subject. However, just like book not read, the computer becomes meaningless as an educational tool unless it is engaging enough for students to actually use it. Student engagement puts the interactive qualities of the system in the forefront, not in terms of superficial "usability", but rather as to "learnability" and "teachability". This poses an array of non-trivial AIED research questions as to how the technology functions in the educational context and how AI may serve to improve the scientific basis of education. Next, I consider some of these challenges in greater detail and how they are dealt with in two types of AIED applications that exemplify the present and future potential of the field.

3 What AIED Brings to Artificial intelligence for Authentic Learning

Education is a social process characterized by learning in interaction: between teachers and students, among students, and, if AIED has a say, between "intelligent" artefacts and both teachers and students. As extensively demonstrated by Reeves and Nass' "media equation" [7], adding interactivity to a system naturally invites social behavior. Accordingly, as computerized learning environments become increasingly interactive and adaptive, they can be said to expand upon the social dimension that affects the learning process. This has important implications for AIED, first for distinguishing AIED applications from other, "static" learning material such as text books; second, for which methods should be used to study learning outcomes (e.g. some would take this as an argument for a situated perspective on learning, or against "media comparison studies" that undervalue the instructional process as such [e.g. 8, 9]). Stressing the interactivity aspect also brings forth the "social intelligence" of the system as an important consideration for new AI techniques.

With the more advanced interactivity that comes with technical development, it makes sense for a field like AIED to take social motivations for learning with artificial systems to the core of its interest. Considering that students may vary as much in what keeps them motivated and engaged as they do in cognitive abilities, AI techniques devoted to exercise social influence (e.g. by virtual agents and feedback) that adapt to the individual student would allow for unique educational arrangements. More specifically, AIED may serve to dissolve previous conceptions of "education", noted also by Cumming and McDougall, as something that takes place in groups (e.g. in schools and classes) versus "learning" as something that happens in an individual (sometimes with a book or a computer).

Schwartz and colleagues [10] argue for a specific form of computer interactivity that generates a "learning sweet spot" of both high motivation and high learning from students' social motivations. This "sweet spot" is achieved through designing an environment that encourages shared initiative and engagement in interaction, in their case with a teachable pedagogical agent. The researchers make the important point that the technology is not used with the goal to "perfectly" model human traits, conversation or intelligence, but only to be *sufficient* to elicit the social schemas (e.g. that of teacher/student) that engage students in productive interactions for learning.

Notably, the research question and goal of making a human-like functioning, artificial system (e.g. "How can we model human intelligence and learning?") are essentially different from those of making a system designed to help students organize and reason with their own concepts (e.g. "How can we visualize students' knowledge?", "What level of prompting from the system is optimal for triggering students to contribute with their own knowledge?"). One can also say that the goal of the former is to produce an *autonomously* intelligent system, whereas the goal of the latter is to produce a *jointly* (with the student and to the advantage of the student) intelligent system.

The perspective on "joint intelligence" takes more of the social context into account, which brings AIED closer to traditional educational research, drawing from established pedagogical strategies such as peer tutoring and learning-by-teaching. However, what makes AIED unique as a research field is that it deals with variables known to have an impact on learning but that have never before been possible to manipulate independently and systematically, such as social roles (including altering avatar representations of gender or ethnicity), others' skills or knowledge level (using virtual peers), and parameters for individually adapting material for teaching in groups. The educational impact of such manipulations makes a prominent subject for AIED research.

In sum, the social nature of interaction points to a range of issues crucial to understanding the impact of AI techniques when employed in real-world educational settings, which AIED should serve to make explicit for both the AI and education community:

First, it is essential to understand which social factors drive students' learning, since technical functions, even when used, may turn out to be used in unintended ways that diverge from the original pedagogical principles [6].

Second, it is important to realize that employing AI techniques may bring "added value" to traditional teaching but possibly also "reduced value", if it takes time and resources from educational needs that are better met by human teachers or other means [11]. It should be a concern of educators to determine when and for what purpose to use AI-based systems for students' learning, and it should be a primary concern of AIED research to distinguish between the "added" and the "reduced" values for different knowledge needs.

Third, and arguably the most crucial point for positioning the AIED research field, the shortcomings of technology, as well as the shortcomings of educational practices for predicting successful learning, leave space for new and original research on what forms students' learning experience in their interaction with technology. I take two examples to show how our expectations of how AI techniques "should" work can be as important for the outcome of the interaction as the underlying technology itself. This also shows that even if the social context cannot be fully controlled or predicted, a careful design can devote AI techniques to create certain "illusions" of intelligent behavior that promote students' learning.

3.1 Example 1: "Educational" games

Many computer games make use of some AI; however, AI techniques appear strikingly underused in the subcategory of so-called "educational games". However, there is an ambiguity in the term "educational games" that both confuses researchers and confounds educational practices. This confusion is mirrored in the debate of whether or not "computer games" as such are effective for learning (for contrasting accounts, see [12, 13]).

In short, the "educational" in games might refer to the educational *subject content* in terms of topics relevant to the curriculum (such as "a math game training fractions"[1] or "a political strategy game that models the global economy"[2]), or it might refer to the (intended) educational *use* of the game in a school context (such as playing a commercial game in the *Halo* or *Assassin's Creed* series that employ AI techniques and that some teachers may use for training "strategy thinking" or "problem-solving" skills, though these are not explicit aims of the game). Games with subject-relevant content typically include explicit exercises or "game tasks" for the intended skill (e.g. counting, spelling tasks) whereas the educational use of other games typically assumes that relevant skills are learned implicitly, through the practice of other kinds of overarching "game goals".

As to the vast offer of games that claim educational content, there is rarely any advanced AI to direct or scaffold the learning process. For example, several AIED-relevant review- and development articles have remarked that the vast majority of math games in the open market (e.g. in AppStore) do not adhere to even basic, cognitive design principles and they seem to contain little more than simple 'drilling' exercises with limited feedback [14, 15, 16].

Using other, commercial computer games for educational purposes, may have great effects on student engagement but little or no effect on learning [13]. This is because game-players may utilize the affordances in a game in a relatively superficial way, learning only "what to press when" to achieve certain results, such that good game performance and progression do not necessarily require the deeper cognitive processing wished for good education. Linderoth reaches the thought-provoking conclusion that the educational appeal of computer games may come from maintaining an "*illusion of learning"* (ibid, p. 59).

The task for the educator is further complicated by the fact that some commercial games might indeed require great skill (e.g. for solving puzzles) but it is hard to determine how much of these abilities are trained by the game itself and, if so, to what extent they transfer to school-relevant tasks (e.g. solving physics problems or mathematical equations).

Nevertheless, it seems safe to say that the education industry has failed to take on board the creative application of the relatively sophisticated "game AI" techniques used in the commercial gaming industry. Rather than the "illusion of learning" on

---

1 http://www.mathsgames.com/fraction-games.html

2 http://www.positech.co.uk/democracy3/index.php

behalf of the student, an alternative and productive use of “game AI” (or “weak AI”) would be to maintain an “illusion of intelligence” [17] on behalf of the software that keeps the student engaged in intellectual activities for actual learning, just like the game-player is kept engaged in playing for entertainment. In other words, a game might not need cognition-like computational power to have educational value in terms of meaningful learning, but the resources it does use should be dedicated to educationally relevant goals. It is to the latter point commercial games often cut short.

Whereas the above may be bad news to educational games, it is good news to the AIED research field, because it shows that there are both *potentially* effective AI techniques already in place and an enormous interest from the education community to employ them (e.g. Education apps being the second largest category in AppStore after Games, both in numbers of 100.000+). Games appear as a domain where AI and Educational interests merge but where AI techniques have been underemployed *for learning*. This makes “educational games” a primary topic for AIED research, a brain child still in its infancy, which calls for more attention and better interdisciplinary upbringing by AI and Ed.

3.2 Example 2: ITS including Teachable Agents

Intelligent Tutoring Systems (ITS) might represent the most dedicated and successful use of AI for student-centered learning, since it actively employs AI techniques not only to structure and present material but also for communication purposes that (more or less) model that of a human teacher. I chose this example (and to include Teachable Agents, or TA, as a “reversed” model of tutoring) because it clearly illuminates how AI-based system can make use of familiar social schemas [e.g. 18, 19].

ITS and TA exploit and benefit from social learning mechanisms (most visibly so when represented by a visual character on screen although the system could be entirely text-based) derived from the student-teacher relationship. As these systems become increasingly advanced, the knowledge needs about people’s social motivations and social psychology in general become of greater importance to the AIED field.

But is it a realistic, or even wanted aspiration, for AIED purposes (i.e. for use in teaching and learning) to develop virtual agents that are as life-like or sociable as a real person? This is an important question for the future of AIED because it poses where resources are better spent; for instance, how should the overwhelming task of producing human-like AI be balanced against working out effective instructional strategies that can be formalized and computed?

Importantly, artificial systems can invoke social responses to improve learning without having to employ AI. For example, Okita et al [20] showed that the mere belief in “real” social interaction when interacting with a computer agent had positive effects on learning, again an effect exploited in the TA metaphor [18]. Some of my own AIED research [3, 4] suggests that social effects of interacting with a Teachable Agents might also transfer from the learning situation to being tested on one’s knowledge; students took on harder problems and performed better on those problems if tested “in company” with the TA they had previously worked with in a learning game. In short, remarkably simple social stimuli may trigger complex and beneficial learning behaviors.

As an interesting contrasting example, some researchers have employed AI techniques to create an "illusion of teachability", by making an agent appear more socially sensitive to the student's input than it actually is [21]. In this case, the system construes a mental model of the human student that informs the agent's responses so it appears as "teachable", though it actually only reflects the kind of knowledge gaps and mistakes that the student has displayed. In effect, the student has to "teach" exactly the things needed to improve his/her own shortcomings (and not necessarily those of a third-party agent). This adds to the power of the social schema by showing that not only the intentional belief of teaching drives the effect, but also the belief in how the tutee (the agent) responds.

My point here is that an important topic of AIED research is to disentangle the social and cognitive mechanisms underlying the effects of ITS and TA, both for the general understanding of such systems and for developing resource-effective systems. For example, the "teachability" features of a TA may be theoretically divided into the underlying (AI-governed) mechanisms that direct the information processing, and its social appearance, as constituted by its visual looks, the things it says, and the types of choices it offers. A key contribution of technology is to offer means to control and regulate these factors through digitalized and personalized "social" responses that can avoid the pitfalls of human socializing (such as distraction from the task and negative stereotyping) while maintaining and even adding to the benefits (such as constructive feedback and active engagement).

In sum, ITS and TA represent a case of true cross-fertilization of the AI and Education domains that produce some unique results, never before seen in human history: semi-independent, virtual beings whose sociable qualities place them somewhere in between artificial and human agents, more like active "educational peers" than passive information systems. In this sense, AIED breaks up the traditional teacher/student dichotomy and includes a third party in the educational design. Students' social motivations to engage in interaction with this party might be more a matter of the effective representation of social features as learned and recognized from the outside world, than how its knowledge is represented inside the system. For use in the social context of a classroom, this makes a strong argument for bringing in more of the educator's experience of "what works" into the design of AI systems.

4 Moving On

Taking the example of cognitive science, I aimed to illustrate AIED as a multidisciplinary practice that forms its identity in relation to technical development as well as pedagogical methods and social learning theories. To the extent that "AI" and "Education" hold separate identities as distinct fields that do not seamlessly combine or "marry", it might be more productive to focus on what they can form together, as a common theme for their future. Educational games and ITS with TA provide example domains that cannot be said to be either "AI" or "Ed" but very much AIED. Relating to those examples, I conclude by suggesting two further discussion points that AIED should take into consideration when moving on together:

1. As to educational software (including games, ITS, simulations and other digital learning environments), AIED still seems predominantly concerned with development and design aspects, whereas little has been done to serve educators' need for sound evaluation and scientifically based, qualitative assessment of existing applications. How do design criteria for learning-effective software translate to evaluation criteria? Considering the vast selection of educational apps to date, perhaps the best way to guide teachers is to formulate meta-criteria that help inform their own selection and recognize well-designed content? How can AIED assist in making this judgment scientifically informed?

2. Considering the range of issues an AIED researcher may have to confront, as exemplified in this text, what is the essence of the "AIED competence" – what does an AIED researcher (need to) know that others don't? Is there anything such as "interdisciplinary expertise" in its own right and then, how does this show, and how is it applied, within AIED research? Is the explication of specific AIED knowledge areas required (or just helpful) for forming a unique identity of the field?

References

1. West, D.: Digital schools: How technology can transform education. Brookings, Washington D.C. (2012)
2. Cumming, G., & McDougall, A.: Mainstreaming AIED into education?. International Journal of Artificial Intelligence in Education (IJAIED), 11, 197-207. (2000)
3. Sjödén, B., Tärning, B., Pareto. L., Gulz, A.: Transferring teaching to testing, An unexplored aspect of teachable agents. In: Proc. of the 15th Int. Conf. on Artificial Intelligence in Education (AIED 2011), LNAI, vol. 6738 (pp. 337-344). Springer, Heidelberg (2011)
4. Sjödén, B., & Gulz, A.: From Learning Companions to Teaching Companions: Experience with a Teachable Agent motivates students' performance on summative tests. In: Proc. of the 17th Int. Conf. on Artificial Intelligence in Education (AIED 2015), *in press*
5. Pareto, L., Haake, M., Lindström, P., Sjödén, B., & Gulz, A.: A teachable agent based game affording collaboration and competition – evaluating math comprehension and motivation. Educational Technology Research and Development, 60(5), 723-751 (2012)
6. Lindström, P., Gulz, A., Haake, M., Sjödén, B.: Matching and mismatching between the pedagogical design principles of a math game and the actual practices of play. Journal of Computer Assisted Learning, 27(1), 90-102 (2011)
7. Reeves, B., & Nass, C.: The Media Equation. How people treat computers, television, and new media like real people and places. CSLI Publications and Cambridge university press. (1996)
8. Clark, R. E.: Reconsidering the research on learning from media. Review of Educational Research, 53(4), 445-459 (1983)
9. Ross, S. M., Morrison, G. R., & Lowther, D. L.: Educational technology research past and present: Balancing rigor and relevance to impact school learning. Contemporary Educational Technology, 1(1), 17-35 (2010)
10. Schwartz, D. L., Blair, K. P., Biswas, G., Leelawong, K., & Davis, J.: Animations of thought: Interactivity in the teachable agent paradigm. In R. Lowe & W. Schnotz (Eds.), Learning with Animation: Research and Implications for Design (pp. 114-140). UK: Cambrige University Press. (2007)

11. Chin, D. B., Dohmen, I. M., Cheng, B. H., Oppezzo, M. A., Chase, C. C., & Schwartz, D. L.: Preparing students for future learning with teachable agents. Educational Technology Research and Development, 58(6), 649-669 (2010)
12. Gee, J. P.: What Video Games Have to Teach Us About Learning and Literacy. Palgrave Macmillan, New York (2003)
13. Linderoth, J.: Why gamers don't learn more: An ecological approach to games as learning environments, Journal of Gaming and Virtual Worlds, 4: 1, pp. 45–62, doi: 10.1386/jgvw.4.1.45_1 (2012)
14. Larkin, K.: Mathematics Education. Is there an App for that?. In: Mathematics education: Yesterday, today, and tomorrow, pp. 426-433. Mathematics Education Research Group of Australasia (MERGA) (2013)
15. Veenstra, B., Van Geert, P. L. C., Van der Meulen, B. F.: Is edutainment software really educational? A feature analysis of Dutch edutainment software for young children. Netherlands Journal of Psychology, 66(2), 50-67 (2011)
16. Ginsburg, H. P., Jamalian, A., Creighan, S.: Cognitive guidelines for the design and evaluation of early mathematics software: The example of MathemAntics. In: Reconceptualizing early mathematics learning (pp. 83-120). Springer Netherlands (2013)
17. Buckland, M.: Programming game AI by example. Jones & Bartlett Learning, Burlington (2005)
18. Chase, C., Chin, D., Oppezzo, M., Schwartz, D.: Teachable agents and the protégé effect: Increasing the effort towards learning. J. of Sci. Edu. and Tech., 18, 334-352 (2009)
19. Ogan, A., Finkelstein, S., Mayfield, E., D'Adamo, C., Matsuda, N., Cassell, J.: Oh dear stacy!: social interaction, elaboration, and learning with teachable agents. In: Proc. of the SIGCHI Conference on Human Factors in Computing Systems, pp. 39-48. ACM (2012)
20. Okita S, Bailenson J, Schwartz D.: The mere belief of social interaction improves learning. In: McNamara DS, Trafton JG (eds) The proceedings of the 29th meeting of the cognitive science society. August, Nashville, pp 1355–1360 (2007)
21. Lenat, D. B., & Durlach, P. J.: Reinforcing Math Knowledge by Immersing Students in a Simulated Learning-by-teaching Experience. International Journal of Artificial Intelligence in Education, 24(3), 216-250. (2014)

AI and Education: Celebrating 30 years of Marriage

Beverly Park Woolf[1]

[1]College of Information and Computer Sciences
University of Massachusetts Amherst, USA

Abstract. This article describes contributions that artificial intelligence (AI) has made and needs to continue to make towards long-term educational goals. The article articulates two challenges in education that require the use of AI: personalizing teaching and learning 21st century skills. This article first describes AI and some of its history and then suggests why AI is invaluable to development of instructional systems. Instructional systems that use AI technology are described, e.g., computational tools that personalize instruction, enhance student experience and supply data for development of novel education theory development. Additionally, some intelligent tutors supply researchers with new opportunities to analyze vast data sets of instructional behavior and learn how students behave.

1 A Brief History of Artificial Intelligence in Education

The field of Artificial Intelligence in Education is focused on research into, development of and evaluation of computer software that improves teaching and learning. Several long term goals have been espoused, such as to interpret complex student responses and learn as they operate; to discern where and why a student's understanding has gone astray, to offer hints to help students understand the material at hand and ultimately to simulate a human tutor's behavior and guidance. Personalized tutors have been envisioned that adapt to an individual student's needs or to teach to groups of students, e.g., classified by gender, achievement level, amount of time for lesson, etc. Another goal is to use Artificial Intelligence (AI) techniques learn about teaching and learning and to contribute to the theory of learning.

AI techniques are needed for almost every phrase in the definition of intelligent tutors above, including *interpret* complex student responses, *learn* as they operate, *discern* where and why a student's understanding has gone astray and *offer* hints. The central problems (or goals) of AI research include reasoning, knowledge, planning, learning, natural language processing (communication), perception and the ability to move and manipulate objects [1]. AIED has been applied to complex domains, e.g. physics, programming, writing essays, and reading. These tutors learn about the strengths and weaknesses of students in these domains and also about students' skills, and emotion. How effective are intelligent tutors? Several tutors have been shown to be very effective in the classroom. Researchers looking at student skills at end of experiments and also at the end of course and large scale standardized testing evaluations found dramatic improvement understanding and learning [2]. Intelligent online tutors are an AI success story [3], though researchers seek to move beyond domain dependence and to support learning of multiple tasks and domains.

To mentor effectively and support individuals or groups, intelligent tutors will assess learning activities and model changes that occur in learners. Estimates of a learner's competence or emotional state, stored in user models, represent what learners know, feel, and can do. When and how was knowledge learned? What pedagogy worked best for this individual student? Machine learning and data mining methods, both derived from the field of AI, are needed to explore the unique types of data that derive from educational settings and use those methods to better understand students and the settings in which they learn (see [2, 4]).

Technology cannot impact education in isolation, rather it operates as one element in a complex adaptive system that considers domain knowledge, pedagogy and environments that students, instructors and technology co-create [5]. AI and Education researchers need to be driven by the problems of education practice as they exist in school settings. The emerging forms of technology described here will challenge, if not threaten, existing educational practices by suggesting new ways to learn [6]. Policy issues that involve social and political considerations, need to be addressed, but are beyond the scope of this document.

2 AI called by a different name: AI behind the scenes

Many components of intelligent instructional systems have their roots in artificial instructional research, e.g., adaptive curriculum, modeling (student, teacher, domain), educational data mining, speech recognition and dialogue systems. All began by using artificial intelligence (AI) techniques. Yet once these algorithms and techniques begin to appear as parts of larger tutors, the tutors are no longer considered AI and AI receives little or no credit for their successes. Many of AI's greatest innovations have been reduced to the status of just another item in the tool chest of instructional designers or computer science. Nick Bostrom explains "A lot of cutting edge AI has filtered into general applications, often without being called AI because once something becomes useful enough and common enough it's not labeled AI anymore." [7] "After all, all smart technologies currently in use (in the classrooms or homes), from tablet computers to smart phones, from Internet search engines to social networking sites, have a growing reliance on techniques derived from AI." [7] The AI effect began in the larger AI field and "occurs when onlookers discount the behavior of an artificial intelligence program by arguing that it is not real intelligence." [7] Pamela McCorduck writes: "It's part of the history of the field of artificial intelligence that every time somebody figured out how to make a computer do something—play good checkers, solve simple but relatively informal problems—there was chorus of critics to say, 'that's not thinking'." [8] AI researcher Rodney Brooks complains "Every time we figure out a piece of it, it stops being magical; we say, 'Oh, that's just a computation.'" [9].

Intelligent personal assistants in classrooms or in smartphones use algorithms that emerged from lengthy AI research. IBM's question answering system, Watson, which defeated the two great Jeopardy champions by a significant margin, was derived from basic AI research in natural language processing, information retrieval, knowledge representation, automated reasoning, and machine learning technologies to the field of open domain question answering [10]. In addition, the Kinect, which provides a 3D body–motion interface for the Xbox 360 and the Xbox One was derived from basic AI research [7].

AI is whatever hasn't been done yet. Software and algorithms developed by AI researchers are now integrated into many applications, without really being called AI, e.g., speech understanding as part of online travel reservations, expert systems that save companies millions of dollars (US). Michael Swaine reports "AI advances are not trumpeted as artificial intelligence so much these days, but are often seen as advances in some other field."[11] "AI has become more important as it has become less conspicuous," Patrick Winston says. "These days, it is hard to find a big system that does not work, in part, because of ideas developed or matured in the AI world." [12].

3 Impact on Education

A related question about AIED relates to the impact of AI on education and focuses on the extent to which the results of AIED research are meaningful to real educational practice [13]. Does the education community even care? Similar to many fields aspiring to scientific rigor, the AIED community can showcase dozens of studies demonstrating the statistical significance of this or that approach or system or their individual components through rigorously designed studies, but it is not always clear how the results of many of those studies actually translate into real educational teaching and learning practices raising a question as to whether all this rigor may not be happening in a vacuum.

For example, schools in the USA are not thriving. Too many schools *teach in traditional ways* and aren't preparing the next generation to meet new challenges. When today's students graduate, they'll be asked to fill the jobs of tomorrow—ones we can't even imagine [14]. And they'll be asked to tackle global problems like climate change, endemic hunger, and refugee problems. Additionally, the current use of digital resources in K12 and higher education can be described as dysfunctional: many school stakeholders can't find sufficient effective digital resources, while large collections of resources exist and sit online, waiting to be discovered. Some solutions have been proposed to migrate successful evidence-based digital resources into classrooms. One solution is to define a roadmap that moves well-tested resources towards publishers and software companies and ultimately into classrooms.

More than 4 million USA students at the K12 level took an online course in 2011, up significantly from just 1 million three years earlier. During the coming decade education should shift from print to digital and from batch processing to personalized learning [15]. In addition to virtual schools, online learning is increasingly being incorporated into traditional settings that blend the best of online and face- to-face learning. A shift to online learning is happening in K12 in the USA due in part to the need to implement college- and career- ready standards, the shift to next-generation assessments, and the prevalence of affordable devices. Online learning may move standardized teaching towards more personalized instruction without increasing the number of teachers.

The field of AIED, now nearly thirty years old, has finally achieved some of its oldest goals. Thirty years is calculated from the first Intelligent Tutoring Systems Conference, 1988, organized by Claude Frasson in Montreal, Canada. Some long-term goals are currently being worked on, including understanding and responding to student knowledge, meta-knowledge (thinking about learning), and affect [16-19]. Educational games and new forms of digital learning are being investigated. In many cases evaluation of student progress shows improvement in learning. Some of the success is due to increasing computer power and some due to researchers focusing on specific isolated problems and pursuing them with the highest standards of scientific accountability. The reputation of AIED, in the education world at least, is still not very positive, because few tutors are robust enough to work consistently in a classroom environment.

4 Future directions for AIED to justify and maintain its unique identity

AI techniques are essential to develop new representations and reasoning about cognitive insights, to provide a rich appreciation of how people learn and to measure collaborative activity. Communities of researchers offer distinct clues to further refine individual instruction in online environments and also require far deeper knowledge about human cognition, including dramatically more effective constructivist and active instructional strategies [20].

4.1 Personalize teaching

One-to-one attention is very important for learning at any age. Research has also shown that students' emotions influence achievement outcomes: confidence, boredom, confusion, stress, and anxiety are all strong predictors of achievement [21, 22]. However, teachers are unable to provide attention based on intimate knowledge of each student. Providing personalized teaching for every learner begins by providing timely and appropriate guidance for student cognition, meta-cognition and emotion [20]. In other words, online tutors should determine in real-time what to say, when to say it, and how to say it. This process grows increasingly complex as the topics become more difficult and the required detectors becomes more complex, e.g., detectors for students' knowledge, skills, or emotion. The field of Learning Science has provided a wealth of knowledge about how to deliver effective feedback and how to teach with new methods (e.g., problem-based learning [23]. Rich, multi-faceted models of instruction go beyond providing simple statements about correctness and provide feedback appropriate to each student's learning needs.

Mentoring systems should support learners with decision-making and reasoning, especially in volatile and rapidly changing environments. Learners often need to make informed decisions and justify them with evidence, gathered through collaboration and communication (see [24, 25]). Students need to learn science practices, scientific reasoning and how to apply facts and skills they have acquired. In collaborative learning, students share their experiences and perhaps persuade others to see their point of view, and articulate what they need to learn more about. They "mess about" and generate their own questions about the targeted science. Groups of students need to be supported as they discuss their methods and results, ask questions and make suggestions.

Respond to student affect. Student emotion while learning is critical to understanding student behavior. Researchers are developing intelligent tutoring systems that interpret and adapt to the different student emotional states [26, 27]. Humans do not just use cognitive processes to learn; they also use affective processes. For example, learners learn better when they have a certain level of disequilibrium (frustration), but not enough to make the learner feel completely overwhelmed [28]. This has motivated researchers in affective computing to produce and creating intelligent tutoring systems that can interpret the affective process of students. An intelligent tutor can be developed to read an individual's expressions and other signs of affect in an attempt to find and guide the student to the optimal affective state for learning. There are many complications in doing this since affect is not expressed in just one way but in multiple ways so that for a tutor to be effective in interpreting affective states it may require a multimodal approach (tone, facial expression, etc.). One example of a tutor that addresses affect is Gaze Tutor that was developed to track students' eye movements and determine whether they are bored or distracted and then the system attempts to reengage the student [29].

AI might be a game changer in education. It provides tools to build computational models of students' skills and to scaffold learning. AI methods can act as catalysts in learning environments to provide knowledge about the domain, student and teaching strategies through the integration of cognitive and emotional modeling, knowledge representation, reasoning, natural language question-answering and machine learning methods [30]. When such tutors work smoothly they provide flexible and adaptive feedback to students, enabling content to be customized to fit personal needs and abilities and to augment a teacher's ability to respond. AI techniques appear to be essential ingredients for achieving mentors for every learner.

User models are being developed that leverage advanced reasoning and inference-making tools from AI, represent inferences about users, including their level of knowledge, misconceptions, goals, plans, preferences, beliefs, and relevant characteristics (stereotypes) along with records of their past interactions with the system. They might also include information on the cultural preferences of learners [31] and their personal interests and learning goals. When modeling groups of learners, the model should make inferences to identify the group skills and behavior.

Finally, providing a mentor for every learning group means improving the ability of intelligent tutors to provide timely and appropriate guidance. In other words, tutors need to determine in real-time what to say, when to say it, and how to say it. This grows more complicated as the skills demanded by society increase in complexity. The learning sciences have provided a wealth of knowledge about how to deliver effective feedback, but the challenge is to incorporate 21st century skills, such as creativity and teamwork.

4.2 Teach 21st Century Skills

Citizens of the 21st century require different skills than did citizens from earlier centuries [20]. 21st century skills include cognitive skills (non-routine problem solving, systems thinking and critical thinking), interpersonal skills (ranging from active listening, to presentation skills, to conflict resolution) and intrapersonal skills (broadly clustered under adaptability and self-management /self-development personal qualities) [32]. We describe two AI techniques that can improve teaching for 21st Century skills: dialogue systems and inquiry learning.

Dialogue Systems. One key development for teaching 21st century skills is implementation of strong dialogue and communication systems. Human tutors can understand a student's tone and inflection within a dialogue and interpret this to provide continual feedback through ongoing dialogue. Intelligent tutoring systems are still limited in dialogue and feedback. Systems that begin to simulate natural conversations have been developed [33, 34]. However, more research is needed to understand student tone, inflection, body language, and facial expression and then to respond to these. Dialogue modules in tutors should ask specific questions to guide students and elicit information while supporting them to construct their own knowledge [33, 34]. The development of more sophisticated dialogues between computers and students partially addresses the current limitations in human-computer communication and creates more constructivist teaching approaches.

The 21st century worker needs both 'hard' skills (traditional domains, such as, history, mathematics, science) as well as 'soft' skills (teamwork, reasoning, disciplined thinking, creativity, social skills, meta-cognitive skills, computer literacy, ability to evaluate and analyze information). Further, working in today's knowledge economy requires a high comfort with uncertainty, a willingness to take calculated risks, and an ability to generate novel solutions to problems that evade rigorous description. Unfortunately, many of today's classrooms look exactly like 19th century classrooms; teachers lecture and students remain passive and work alone on homework problems that do not require deep understanding or the application of concepts to realistic problems. Our system of education is behind and the gap grows wider each day.

As we know, changes in educational policy, practice and administration tend to happen slowly. For example, in the U.S. about 25 years are required for an individual to receive a sufficiently well-rounded education to become a proficient educator [30, 35]. The impact of that individual's teaching cannot be seen in subsequent learners for another 20 years. Thus the total cycle time for learning improvement is on the order of 45 to 50 years. Very few challenges in research or social policy cover such a long time scale [36].

Inquiry and Collaborative Learning. What type of technology is needed to mentor students as they learn complex, ill-structured problems? How can technology support exploratory behavior and creativity? Open-ended and exploratory inquiry-based systems support learners to question and enhance their understanding about new areas of knowledge [37, 38]. Innovative instructional approaches, such as preparation for future learning, have uncovered ways to increase comfort with uncertainty and promote development of adaptive expertise [39].

Engagement in the information society often requires people to collaborate and exchange real-time responses over lengthy time periods [20]. A single individual working alone over time often cannot provide enough expertise to solve modern problems (e.g., environmental issues, sustainability, security). Technology is needed to support small groups, class discussions, 'white boarding,' and the generation of questions. To support learners in groups, networking tools are needed to facilitate individuals to learn within communities, communities to construct knowledge, and communities to learn from one another [40-43]. AI software is needed to support students in collaboration, researchers to examine learning communities and learning communities to morph into global communities. For example, how do learning communities sustain, build on, and share knowledge? Students clearly do not construct original knowledge in the same way as do research communities, but they can learn from community-based project work [44].

Support for inquiry and collaboration is needed as students become exposed to diverse cultures and viewpoints. What is the process by which teams generate, evaluate, and revise knowledge? How can we enhance learners' communication skills and creative abilities? Which tools match learners with other learners and/or mentors taking into account learner interests? Finally research is needed to support exploratory, social, and ubiquitous learning. How can software both support collaboration and coach about content? Can technology support continuous learning by groups of learners in ways that enable students to communicate what they are working on and receive help as needed? Learning communities, networking, collaboration software and mobile and ubiquitous computing are being used to create seamless social learning [41]. Socially embedded and social driven learning is pervasive.

In a society built on knowledge, citizens need to acquire new knowledge quickly, to explore alternative problem solving approaches regularly and to form new learning communities effectively [20]. People need to tackle knowledge challenges and opportunities. For educators, this requires rapid revision of what is taught and how it is presented to take advantage of evolving knowledge in a field where technology changes every few years. As an example of rapid change and unpredictability, consider the Internet itself. It first appeared in the mid-1990s. By 2015, 37.3% of the Earth's population uses it. Internet services and applications apply to virtually every aspect of modern human life (e.g., research, banking, shopping, meeting people, health, travel, job seeking). How can education prepare students for a society that changes so dramatically and rapidly? In just 25 years the Internet has become a major factor in nearly every civilized activity and applies to virtually every aspect of human life. At the minimum, students need to be taught how to search it, learn from it, evaluate its information, use it wisely, and contribute to it with well-vetted information. One answer lies in improved and expanded learner competencies. Learners must be more creative, more agile, and more able to learn in groups; they must know how to learn. Key features include skills in critical thinking, creativity, collaboration, meta-cognition and motivation.

5 Discussion

This article described why AI is vital in Education and identified two challenges: personalized teaching and learning 21st century skills. Specifically, personalized learning should be supported by tools that enhance student and group experience, reflection, analysis, and theory development. Learning 21st century skills should be facilitated by resources that improve human-computer interfaces (dialogue systems) and inquiry-based and collaborative learning. We also expect AI technology to contribute to richer experiences for learners who will then be able to reflect on their own learning. Learning scientists with AI tools will have new opportunities to analyze vast data sets of instructional behavior collected from rich databases, containing elements of learning, affect, motivation, and social interaction.

Research shows that skilled workers have more job opportunities than do less skilled workers [45]. As technology advances, educated workers tend to benefit more, and workers with less education tend to have their jobs automated.

Over the next few years we expect intelligent online instruction to increasingly be a part of the online learning landscape [46]. Maybe in five years, children will increasingly be online with educational games and simulation environments; behind the scene will be intelligent tutoring capabilities adapting the environment. Similar to working with Google, people may not know what the adaptation algorithm is doing, but it is changing the individual search ranking in the background [46]. Algorithms are there and making search more effective. Similarly, students will see action like this in the educational material they use, with intelligence in the background. Intelligent tutors may provide many of the benefits of a human tutor and also provide real-time data to instructors and developers looking to refine teaching methods.

References

1. Russell, S.J. and P. Norvig, *Artificial Intelligence: A Modern Approach*. 3rd ed. 2010, Upper Saddle River: Prentice Hall.
2. Koedinger, K., et al., *New potentials for data-driven intelligent tutoring system development and optimization*, in *AI Magazine, Special Issue on Intelligent Learning Technologies*. 2013.
3. Anderson, J.R., et al., *Cognitive tutors: Lessons learned.* The Journal of the Learning Sciences, 1995. **4**(2): p. 167-207.
4. Conati, C. and S. Kardan, *Student modeling: Supporting personalized instruction, from problem solving to exploratory open ended activities.* AI Magazine, 2013. **34**(3): p. 13-26.
5. Oblinger, D.G., *Game Changers: Education and Information Technology*. 2012, Educause: Washington, D.C.
6. McArthur, D., Lewis, M. and M. Bishay, *The Roles of Artificial Intelligence in Education: Current Progress and Future Prospects*. 1994: RAND DRU-472-NSF.
7. AIE. *AI Effect*. 2015; Available from: Retrieved from http://en.wikipedia.org/wiki/AI_effect.

8. McCorduck, P., *Machines Who Think*. 2nd Edition ed. 2004, Natick, MA: A. K. Peters, Ltd.
9. Kahn, J., *It's Alive*, in *Wired*. 2002.
10. Watson. 2015; Available from: retrieved from http://en.wikipedia.org/wiki/Watson_(computer).
11. Swaine, M. *AI: It's OK Again!, Dr. Dobbs, The World of Software Development*. 2007 May 20, 2015]; Available from: Retrieved from http://www.drdobbs.com/architecture-and-design/ai-its-ok-again/201804174.
12. Hofstadter, D., *Gödel, Escher, Bach: an Eternal Golden Braid*. 1980: Basic Books.
13. AIED. *Call for workshop proposals AIED*. 2015; Available from: Retrieved from http://WWW.sussex.ac.uk/Users/bend/aied2015/.
14. Hewlett, W., *Programs in deeper learning*. 2013, William and Flora Hewlett Foundation.
15. VanderArk, T. and C. Schneider, *How Digital Learning Contributes to Deeper Learning*. 2012, Getting Smart: http://gettingsmart.com/wp-content/uploads/2012/12/Digital-Learning-Deeper-Learning-Full-White-Paper.pdf. p. 26.
16. Arroyo, I., et al., *Emotion sensors go to school*, in *Proceedings of the 14th International Conference on Artificial Intelligence in Education (AIED'09)*, V. Dimitrova, et al., Editors. 2009, IOS Press: Brighton, UK. p. 17-24.
17. Chaouachi, M., et al., *Affect and mental engagement: Towards adaptability to intelligent systems*, in *The 23th International FLAIRS Conference*. 2010, AAAI Press: Daytona Beach, FL, USA.
18. D'Mello, S., A. Graesser, and R.W. Picard, *Toward an affect-sensitive AutoTutor.* Intelligent Systems, IEEE, 2007. **22**(4): p. 53-61.
19. Frasson, C. and A. Heraz, *Emotional Learning. Encyclopedia of the Sciences of Learning*. 2011: Springer Verlag.
20. Woolf, B.P., et al., *AI Grand Challenges for Education.* AI Magazine, 2013. **34**(4): p. 66-83.
21. Goleman, D., *Emotional Intelligence: why it can matter more than IQ*. Bloomsbury. 1996, London.
22. Pekrun, R., et al., *Boredom in achievement settings: Exploring control–value antecedents and performance outcomes of a neglected emotion.* Journal of Educational Psychology, 2010. **102**(3): p. 531-549.
23. Hmelo, C.E., *Problem-based learning: Effects on the early acquisition of cognitive skill in medicine.* The Journal of the Learning Sciences, 1998. **7**(2): p. 173-208.
24. Rus, V., et al., *Recent Advances in Conversational Intelligent Tutoring Systems*, in *AI Magazine, Special Issue on Intelligent Learning Technologies*. 2013.
25. Swartout, W., et al., *Virtual humans for learning*, in *AI Magazine, Special Issue on Intelligent Learning Technologies*. 2013.
26. D'Mello, S.K. and A. Graesser, *Language and discourse are powerful signals of student emotions during tutoring.* IEEE Transactions on Learning Technologies, 2012. **5**(4): p. 304-317.
27. Woolf, B., et al., *The effect of motivational learning companions on low achieving students and students with disabilities.*, in *10th International Conference on Intelligent Tutoring Systems (ITS'10)*, V. Aleven, J. Kay, and J. Mostow, Editors. 2010, Springer Berlin/Heidelberg. p. 327-337.
28. Graesser, A.C., et al., *Detection of emotions during learning with AutoTutor*, in *Proceedings of the 28th Annual Meeting of the Cognitive Science Society*. 2006, Cognitive Science Society. p. 285-290.
29. D'Mello, S., et al., *Gaze tutor: A gaze-reactive intelligent tutoring system.* International Journal of human-computer studies, 2012. **70**(5): p. 377-398.
30. Woolf, B.P., *Building intelligent interactive tutors: Student-centered strategies for revolutionizing e-learning*. 2009, San Francisco, CA: Morgan Kauffman.

31. Blanchard, E.G. and D. Allard, *Handbook of Research on Culturally-Aware Information Technology: Perspectives and Models*. 2010, Hershey, PA: Information Science Publishing.
32. Pellegrino, J. and M. Hilton, Editors, *Committee on Defining Deeper Learning and 21st Century Skills*. 2012, National Research Council: http://www.leg.state.vt.us/WorkGroups/EdOp/Education for Life and Work- National Academy of Sciences.pdf.
33. Graesser, A.C., VanLehn, K. , Rosé, C.P., Jordan, P.W., and Harter, D., *Intelligent tutoring systems with conversational dialogue*, in *AI Magazine*. 2001, American Association for Artificial Intelligence. p. 39-51.
34. Rosé, C., et al., *Analyzing collaborative learning processes automatically: Exploiting the advances of computational linguistics in computer-supported collaborative learning.* International journal of computer-supported collaborative learning, 2008. **3**(3): p. 237-271.
35. King, J., N. Sabelli, and H. Kelly, *Preamble on Policy Issues*, in *Global Resources for Online Education (GROE) Workshop*. 2009: Tempe, AZ.
36. Roschelle, J., et al., *Eight issues for learning scientists about education and the economy.* Journal of the Learning Sciences, 2011. **20**(1): p. 3-49.
37. Dragon, T., B.P. Woolf, and T. Murray, *Intelligent coaching for collaboration in ill-defined domains*, in *Proceedings of the 2009 conference on Artificial Intelligence in Education: Building Learning Systems that Care: From Knowledge Representation to Affective Modelling* 2009, IOS Press. p. 740-742.
38. Floryan, M. and B.P. Woolf, *Improving the efficiency of automatic knowledge generation through games and simulations*, in *Proceedings of 16th International Conference on Artificial Intelligence in Education (AIED-2013). LNAI 7926*, K. Yacef, Editor. 2013, Springer: Heidelberg.
39. Schwartz, D.L. and T. Martin, *Inventing to Prepare for Future Learning: The Hidden Efficiency of Encouraging Original Student Production in Statistics Instruction.* Cognition and Instruction, 2004. **22**(2): p. 129-184.
40. Suthers, D., *Representational support for collaborative inquiry*, in *System Sciences, 1999. HICSS-32. Proceedings of the 32nd Annual Hawaii InternationalConference on System Sciences* 1999, IEEE. p. 14-.
41. Suthers, D.D., *Representational guidance for collaborative inquiry*, in *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments.*, J. Andriessen, M.J. Baker, and D.D. Suthers, Editors. 2003, Kluwer Academic: Dordrecht. p. 27-46.
42. Suthers, D.D. and C. Hundhausen, *An experimental study of the effects of representational guidance on collaborative learning processes.* Journal of the Learning Sciences, 2003. **12**(2): p. 183-219.
43. Woolf , B.P., *A Roadmap for Education Technology.* 2010, A Report to the Computing Community Consortium: http://telearn.archives-ouvertes.fr/docs/00/58/82/91/PDF/groe_roadmap_for_education_technology_final_report_003036v1_.pdf. p. 80 pp.
44. Johnson, D.W., Johnson, R., *An Overview of Cooperative Learning*, in *Creativity and Collaborative Learning*, J.Thousand, A. Villa, and A. Nevin, Editors. 1994, Brookes Press: Baltimore, MD.
45. Brynjolfsson, E. and A. McAfee, *Race Against The Machine: How the Digital Revolution is Accelerating Innovation, Driving Productivity, and Irreversibly Transforming Employment and the Economy*. 2013, Lexington, MA: Digital Frontier Press.
46. Koedinger, K., *7 Things You Should Know About Intelligent Tutoring Systems*, in *EDUCAUSE Learning Initiative (ELI)*. 2013.

AI and Ed: a Happy Open Marriage

Julita Vassileva[1] James Lester[2], Judith Masthoff[3]

[1]University of Saskatchewan, Canada, jiv@cs.usask.ca
[2]North Carolina State University, USA, lester@ncsu.edu
[3]University of Aberdeen, Scotland, j.masthoff@abdn.ac.uk

Abstract. We claim that this marriage has never been closed and exclusive. It started because both AI and Education share the goal of understanding the human process of knowing, and getting to know, i.e. learning. The difference is how the two areas exploit the understanding they aim to develop. AI is more focused on making machines that know and learn like people or better than them. AIED is more interested in supporting people to learn better.

Keywords: AI in Education, AIED, AI

1 Introduction

AI originated from the curiosity of understanding how the human mind works and creating models of reasoning and machines that mimic and improve on human reasoning (using the capacities of computers). The early research in AI started with theoretical studies in reasoning and knowledge representation, metacognition, and learning of a single human (single agent). This research, married the area of Cognitive Psychology and lead to the creation of the area of Cognitive Modeling. The need for practical applications drove the formation of many "children" areas of applied AI: Expert Systems, Probabilistic Reasoning, User Modeling, Ontologies (and more recently, Semantic Web and Linked Data), and Advanced Learning Algorithms (which branched more recently into Data Mining, Data Analytics, Data-warehousing etc.).

Around the mid 1990ies, the theoretical interest shifted towards situated action and social reasoning, and multi-agent architectures, leading to the creation of the area of Multi-Agent Systems. Theoretical studies in Argumentation and Negotiation followed with the creation of their own research areas. The area of Interactive Virtual Agents (IVA) emerged around the end of the 1990ies. Another "child" area of applied AI is Recommender Systems (RecSys), which deploys user modeling and advanced learning algorithms to emerging CS application areas, such as e-commerce. Around the same time, some AI researchers turned their sight to modeling other human psychological phenomena such as emotion and affect, which lead to the establishment of the Affective Computing area.

2 AI and CogPsy Meet Education

AI in Education has been "married" to all of these children of AI. Early ITS work in the 1980s and early 1990ies on pedagogical planning, domain knowledge modeling, student modeling and ITS shells applied techniques from the areas of planning expert systems, and knowledge representation. The second half of the 1990s saw attention shift to agent-based tutoring systems, tutorial dialogues, animated characters, and the first works on modeling learner affect and adapting the interaction with the tutor. In the beginning of the new century the application of ontologies and semantic web technologies for learning material annotation and concept maps for domain knowledge representation took a center stage and the first applications of recommender systems for learning materials considering both content based and social recommendations, and visualizations appeared to explain both the recommendations and the student model (social navigation, open learner modeling). We have seen many research topics in AIED evolve into its own children areas, such as CSCL (a child of AIED, the Learning Sciences and CSCW) and EDM (a child of Data Mining and AIED).

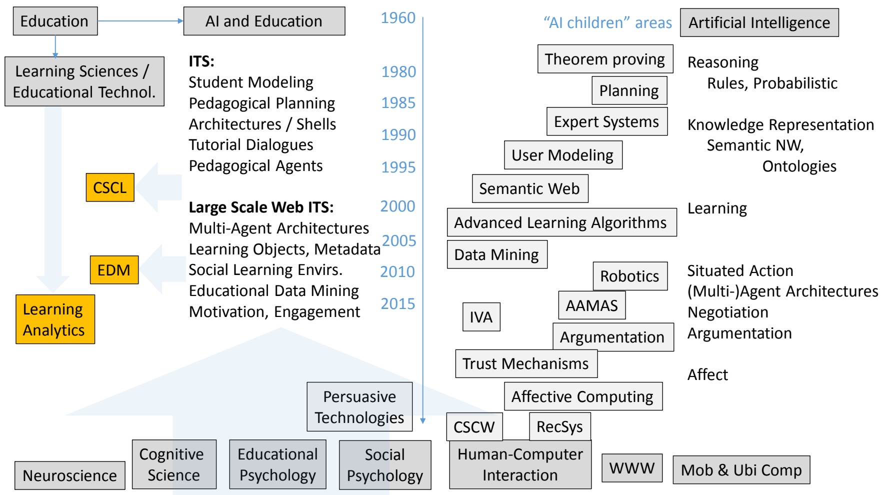


Figure: Approximate evolution of AIED research topics along with AI research topics and the emerging applied AI - children areas, and the influences of other areas of CS and other disciplines AI research topics and

Motivation is an important factor in learning, and the first attempts to model computationally motivational pedagogical strategies started around 1995. The OCC model of emotions triggered interest in incorporating affective factors in HCI around 2000, and it was very soon followed by work in the AIED area, on modeling affect in learning scenarios.

The realization that students engage in off-task behaviours or "game the system" around 2005 lead to increased interest in learner motivation and engagement, as well as educational games (and gamification) which started 10 years earlier. Yet the inspiration for this work is found often in other disciplines (Social Psychology, Persuasive Technology, Behaviour change and even Neuroscience), rather than in Affective Computing.

3. Exploration vs Rigour

The influence of the above AI-children areas, broader computer science areas and other disciplines has been not just in the choice of research topics of AIED researchers, but also in the methods used to carry out, design and evaluate the research. In the early years the focus was on constructing working ITS and a typical research paper included a detailed design justification, description and perhaps a couple of screen-shots as "proof of existence", with not much evaluation. Later on it became necessary to present evaluation data – even if it consisted only of the number of students who liked the system. With the increasing influence of Educational Psychology, evaluation methods from the behavioural sciences were introduced in the area. This coincided with the rapid development of web technologies and tools that allowed an easy design of systems and easier experimentation with more subjects as the ITS prototypes were now accessible on the web. The CMU cognitive tutors were successfully applied with thousands of children in the US, and they started producing a lot of data allowing to evaluate the learning effects on a large scale and for long term use. After 2005, statistical methods for evaluation became a standard, and a typical research paper in the area became much more like a psychology paper or a natural science paper than an engineering paper. The main point became studying the phenomenon of a human interacting with an "experimental tool" designed based on a particular theoretical foundation, and in a way, a significant part of the research in AIED became a branch of applied Cognitive Science. Researchers who were more interested in building systems than in studying human cognition wandered off to other areas that focused more on the technologies, for example ICALT, EC-Tel, Web-Based Learning.

Yet, there are still researchers interested in developing further the "tools", not only from the point of view of the underlying cognitive theories, but also, from the available new technologies developed in the meantime by the Mobile & Ubiquitous Computing community, new data-mining techniques that can allow to automatically learn and improve pedagogical decisions (not necessarily based on theory). The AIED community needs the researchers interested in technology so that the field doesn't become stagnant, overly constrained by methodology, making miniscule improvements based on the same mature AI technologies. So the marriage between AIED and the younger AI children (such as Recommender Systems, or Affective Computing, and even "embryo" areas such as Mechanism Design, Trust, and Negotiation in AAMAS) is important.

When we look at the complex map of how the research topics and children areas emerged, we can notice that in several areas the connection is bi-directional. For example, the area of User Modeling and Personalization, which emerged as a child of AI, has been strongly influenced by AIED. Similarly, IVA, and Affective Computing, have moved ahead to a large extent due to insights and case studies in the context of educational applications. Newly emerged areas, such as Persuasive Technologies also have a lot to learn from the area of AIED, and AIED has a lot to learn from them.

3 Conclusion

So, in conclusion, the marriage between AI and Education and AI is in name only, as much as the name AI describes the inspiration of understanding how the human mind works and creating models (with practical use) of human mind. In fact it has been more of an "open marriage" with quite a few partners – the children areas of applied AI and some other areas and disciplines (as shown in The Figure). Yet, AI is a good family to be in – a large and productive family of smart people. In many ways, it is a perfect marriage.

AI *in* Education as a methodology for enabling educational evidence-based practice

Kaśka Porayska-Pomsta*

University College London, Institute of Education, London Knowledge Lab,
23-29 Emerald Street, London WC1N 3QS K.Porayska-Pomsta@ioe.ac.uk

Abstract. Evidence based practice (EBP) is of critical importance in Education where, increasingly, emphasis is placed on the need to equip teachers with an ability to independently generate evidence of their best practices *in situ*. Such contextualised evidence is seen as the key to informing educational practices more generally. One of the key challenges related to EBP lies in the paucity of methods that would allow educational practitioners to generate evidence of their practices at a low-level of detail in a way that is inspectable and reproducible by others. This position paper focuses on the utility and relevance of AI methods of knowledge elicitation and knowledge representation as a means for supporting educational evidence-based practices through *action research*. AI offers methods whose service extends beyond building of ILEs and into real-world teaching practices, whereby teachers can acquire and apply computational design thinking needed to generate the evidence of interest. This opens a new dimension for AIEd as a field, i.e. one that demonstrates explicitly the continuing pertinence and a maturing reciprocity of the relationship between AI and Education.

1 Introduction

AI methods of knowledge representation and knowledge elicitation can make an important contribution to supporting educational evidence-based practices (EBP) through Action Research (AR). EBP is of critical importance in education where, increasingly, emphasis is placed on the need to equip teachers with an ability to independently generate evidence of their best practices *in situ* [8]. Such evidence is seen as the key to informing educational practices more generally. One of the key challenges related to EBP lies in the lack of readily available methods that would support the generation of evidence by practitioners at a fine-grained level of detail and in a way that is reproducible by other practitioners. There is also a notable lack of consensus as to what constitutes good evidence

* My colleagues Manolis Mavrikis, Karen Guldberg, Sarah Parsons, Helen Pain and Mina Vasalou have all contributed over the years in different ways to the development of the position presented in this paper, as have all of my students taking the *Learning and Teaching with Technologies* module at the UCL Institute of Education. LeAM and TARDIS were both funded by the European Commission (FP6-IST-2003-507826 and FP7-ICT2011-7-288578 respectively).

in education, with randomised controlled studies being typically favoured due to being seen as leading to measurable results similar to those in the biological and medical sciences – currently the gold standard of scientific rigour. Unfortunately, given the inextricable dependency of educational outcomes on the context within which learning and teaching takes place, e.g. [1], the results of such studies tend to have limited generalisability. Education requires a more nuanced and transparent approach than a pill-like medical intervention approaches can offer; they need to serve as tools for teacher reflection and experimentation in order to provide an informed basis for effecting positive change on the learners.

2 In pursuit of a broader definition of AI in Education

AI methods used to elicit knowledge of teaching and learning processes and to represent such knowledge computationally, offer the tools needed by teachers to gather evidence in a systematic, detailed and incremental manner that can be also shared with and inspected by others. Viewing the contribution of AI to Education as a methodological one opens up an important perspective on the possible role of AI in Education than has been adopted to date. Some important fundaments for the adoption of such a perspective have been laid some thirty years ago by Alan Bundy who categorised Artificial Intelligence (AI) field in terms of three kinds of AI: (i) basic AI, aiming to explore computational techniques to simulate intelligent behaviour, (ii) applied AI, concerned with using existing AI techniques to build products for real-world use and (iii) cognitive science, or computational psychology, focusing on the study of human or animal intelligence through computational means [2]. In doing so, Bundy highlighted the diversity of motivations for *doing* AI and, consequently, of the methodologies to both inform and evaluate systems that are underpinned with AI. This diversity of motivations was also noted by Mark and Greer [10] in their exploration of the AIEd evaluations methodologies, where they highlighted the distinction between formative and summative evaluations. Retrospectively, this distinction remains crucial insofar as it allows for a more precise definition of AIEd within the wider fields of AI and Education, by bringing to the fore the dependency between the technologies engineered within AIEd and the purpose, context and design of their use. Over the years, the role of formative evaluation has been elaborated by AIEd researchers based on the growing aspirations of the community not only to establish some ground truths to inform the design and implementation of AIEd technologies, but also to connect AIEd research with educational practices.

Conlon and Pain [5], who relied on Bundy's 3-kind definition of AI to provide their own vision of AIEd, proposed a Persistent Collaboration Methodology (PCM) as a means of ensuring the real-world relevance and effectiveness of the AIEd technologies and to enhance rigour of the design, implementation and evaluation process. PCM draws equally from the key educational methodology of Action Research (AR) [4], applied AI approaches to knowledge elicitation and representation, and human-computer interaction (HCI) design. In contrast with the prevalent practices at the time, PCM advocated that early and continuous involvement of practitioners specifically as *action researchers* in the design and evaluation of AIEd technologies is essential to securing the educational validity of such technologies, to enabling a contribution to both AI and educational theories and practices, and to achieving a balance in the emerging technologies and research between the 'technological push' and 'educational pull'. While inspirational in its effort to acknowledge and marry educational and AI methods PCM remains firmly within the boundaries of AIEd practices offering insights as to the best educational systems designs, but not necessarily as to the best educational practices more generally. In the next two sections I discuss the affordances of knowledge representation as a conceptual tool of relevance to educational practices and, using two examples, I illustrate the role of knowledge elicitation as a means for utilising and for developing this conceptual tool further.

3 Knowledge Representation

Knowledge representation (KR) is fundamental to AI and, arguably, to any scientific endeavour, because at its very basic (and most general), it is a *conceptual* tool for describing and reasoning about the world we inhabit. Scientific theories are in essence forms of knowledge representation about the world, albeit delivered at different levels of specificity. In AI, knowledge representation is inevitably and by definition a theory of intelligence, or more precisely – of intelligent reasoning.

Davis et al. [6] define knowledge representation in terms of five distinct roles that it plays in AI. The first and overarching role of KR, is to serve as a *surrogate* of the thing itself, i.e. the world being represented. As a surrogate, KR offers us (or a computer system) a means for reasoning about the world without having to take action in it, i.e. it allows us to determine consequences within the world we describe by thinking about them rather than by enacting them. Thus, KR provides tools for thinking about and for refining our perceptions of the world, which are, at least conceptual and, at their most usable, computational in nature.

The second role of KR is in forcing us to make *ontological commitments* that tell us how to see the world, i.e. what kind of concepts, entities, etc. and relationships between them describe the world. Since it is impractical (and impossible) to represent all of the characteristics of the world, Davis et al. refer to these ontological commitments as a "strong pair of glasses that determine what we can see, bringing some parts of the world into sharp focus, at the expense of blurring other parts.". They highlight that such focusing/blurring is the greatest affordance of KR in that it enables decisions about what to attend to and what to ignore in our world (Davis et al., [6], p.5). Although ontologies are language agnostic, the choice of representation technologies[1] will impact on what specific commitments we make; logic, rules, frames, semantic nets, etc., constitute different representation technologies, each encapsulating a specific viewpoint on what kinds of things are important in the world. For example, frames use a prototypes viewpoint, whereas logic focuses on individual entities and the relations between

[1] This is the term is used by Davis et al. to refer to "the familiar set of basic representation tools like logic, rules, frames, semantic nets, etc." (p.3)

them. These are by no means the only representation technologies available in AI and neither are they the only technologies that are possible or needed for some domains. In Education and AIEd, ontologies are relatively well understood and accepted as forms of representations of specific subject domains and of knowledge about the learner. However, while they inform us about a possible view of the world, in terms of its component parts, they do not tell us how we can reason about the world using those parts.

The third role of KR is therefore as *a theory of intelligent reasoning*, which tells us what inferences we can and should draw (*sanctions* vs. *recommendations*, respectively), given our ontological commitments. Recommendations define what inferences are appropriate to make and hence which ones are intelligent. A theory of intelligent reasoning lies at the core of AI and, arguably, of educational *practice*, because it is critically concerned with understanding intelligent action and its relationship to the external world [7];[1]. It is this relationship that resides at the heart of teachers' adaptive capabilities and it is in capturing it that one of the greatest challenges for AIEd (*and* Education) lies. This challenge is all the more, because KR related to reasoning involves making the fundamental choice of a theory of intelligent reasoning that must underpin a given representation. Given many different conceptions of intelligent reasoning (e.g. logic, psychology, biology, statistics and economics, etc.) such choice will yield very different conclusions and hence, yet again, different views of the world. For example, logic views reasoning as a form of calculation such as deduction, whereas a theory derived from psychology views intelligent reasoning as a variety of human behaviour, plausibly involving structures such as goals, plans or expectations. Education too offers a variety of different theories of learning, each engendering inferences that are possible and needed. The contrast between approaches which view learning as an outcome of a pre-designed intervention or as an outcome of a transactional experience offers one example.

The fourth role of KR is as a *medium for pragmatically efficient computation*. As such KR provides an environment in which thinking can be accomplished (and conclusions drawn). Ontological and inferential representations jointly provide a contribution to defining such an environment and although they do not in themselves guarantee full computational efficiency, the choice of the specific representation technologies and of intelligent reasoning theory must act in support of achieving such efficiency. While educational theories of learning as transactional and situated experiences are abound they tend to lack specificity as to how exactly such experiences can be captured, described and reasoned about. And while AIEd research provides numerous accounts of such mechanisms and explicitly considers computational efficiency (both as relate to problem solving and affect, e.g. [11]), those accounts tend to be limited in scope and in their power to convince educational community of their applicability to wider education.

The fifth (and final) role of KR is as a *medium of human expression*, i.e. a language through which we convey and ground our view of the world. As such KR allows us to share the different representations with other people. It is precisely the affordance of being sharable and inspectable that makes KR such a compelling candidate as a conceptual tool for supporting evidence-based practices in education. This affordance is also of crucial relevance to AIEd practices: at least in principle, the representations created by educational practitioners can provide rich source of authentic data that can then be used to inform the AIEd systems. However, how successfully the affordances of KR as a medium for expression can *actually* be exploited at the intersection of AIEd and Education, hangs on an understanding that although it does not matter what language we employ to express our world view, the language that we do employ has to be easy to use. As Davis et al. put it "If the representation makes things possible but not easy, then as real users we may never know whether we have misunderstood the representation and just do not know how to use it, or it truly cannot express things we would like to say". Thus, a representation has to provide a language in which we can communicate without having to make a *heroic effort* (p.15).

Davis et al.'s definition of KR in AI is very useful in highlighting its role as a tool for thinking with and as a method for understanding the complexities of our internal and external experiences. There are at least four different ways in which KR as a methodology can serve education. First, it forces us to make explicit our tacit knowledge about the world and the relationships therein. Representing such tacit knowledge enables us not only to reflect on the world that we represent, but also to gain a better understanding of what it is that we actually know. Such reflection is key to educational practice because it brings into focus the strengths and weaknesses in the particular approaches to supporting learning and the kinds of priorities that may characterise such support. Second, KR allows us to create different knowledge representations of the same phenomenon without having to fundamentally change the way we act in the real world. This is important in education where any efforts to effectuate a change involve real and potentially life long impact on real people (the learners) and wherefore such efforts must always be based on informed choices. Third, KR allows us to observe the possible consequences of the different representations on the world, thus enhancing our predictive powers, without involving the actual experience of such consequences. As with the second point, this is important to our being granted access to different viewpoints on the same phenomenon, but this time we also have access to various possible consequences of adopting the different viewpoints. Fourth, KR allows us to share the different representations with other people to generate rich critiques of the different viewpoints and to enrich, update or change our existing viewpoints based on the perspectives of the others' unique experiences and understandings. As well as being shareable with others, KR can also provide a trace of our own views of the world over time and a basis for reflection and introspection on how our ideas evolved and what influenced them.

4 Knowledge elicitation

Knowledge elicitation (KE) is an inseparable companion of knowledge representation in that it is through KE that we engage in reflection about the world.

KE is a *process* in which we can engage alone (through self questioning) or with others, either collaboratively or as respondents to someone else's queries and the process can be either formal or informal, and structured or unstructured.

There are various forms of KE instruments that have been adopted, developed and tested in the context of AIEd. For example, questionnaires or interviews, have been borrowed directly from the social sciences, whereas methods such as post-hoc cognitive walkthroughs, gained in power and applicability with the advent of audio and video technologies, and further through logs of man-machine interactions. Other methods, e.g. Wizard of Oz (WoZ), have been devised as placeholders for yet-to-be-developed fully functional learning environments or components thereof, with the specific purpose of informing the design of technologies in a situated fine-grained level of detail way (e.g. see [12]).

Although KE is standardly employed in AIEd to inform the design of its technologies, its role as a means of explicitly informing educational practice is less well understood and it may be even regarded as somewhat out of AIEd's focus. Yet, it is precisely in examining both how KE informs the design of our technologies and how real educational practices may be affected by KE, that the idea of AI as a methodology, comes to life. It is through this two-way lens that we can start to appreciate the real value of creating a more transitive relationship between AI and Educational practices. Two research projects – LeActiveMath (in short *LeAM*[13]) and TARDIS [14] – serve to illustrate these points.

LeAM is a system in which learners at different stages in their education can engage with mathematical problems through natural language dialogue. It consists of a learner model, a tutorial component, an exercise repository, a domain reasoner and natural language dialogue capabilities. LeAM's design is based on the premise that the specific context of a situation along with the learner-teacher interaction are integral to both regulating learners emotions and to being able to recognise and act on them in pedagogically viable ways.

To inform the learner and the natural language dialogue models, studies were conducted using WoZ design and a bespoke chat interface. Specifically, the student-teacher communication channel was restricted to a typed interface with no visual or audio inputs to resemble the interface of the final learning environment. Five experienced tutors participated in the studies where they had to tutor individual learners in real time, delivering natural language feedback. They were asked to talk aloud about their feedback decisions as they engaged in tutoring and to further qualify those decisions by selecting situational factors, e.g. student confidence or difficulty of material, that they considered important in those decisions. The tutors were asked to make their factor selections through a purpose-built tool every time they provided feedback. To aid them in this task some factors were predefined (based on previous research), but these were not mandatory as the tutors could add their own factors to the existing set.

Following each completed interaction, the tutors were invited to participate in post-task walkthroughs, which synchronised a replay of (1) the recording of the student screen (2) the verbal protocol of the tutor and (3) the selected situational factors for the given interaction. Walkthroughs allowed the tutors and the researchers to review specific interactions, to discuss them in detail, to explain their in-the-moment choices of factors, and to indicate any change in their assessment of the situations.

The data elicited provided a concrete basis for the implementation of LeAM's user and dialogue models and the corresponding knowledge representations. However, the studies also provided important insights into the potential impact that the KE process had on the participating tutors. Specifically, the demand on teachers' to report on the situational factors of importance to their feedback decisions brought to their attention that such factors may indeed play a role and forced them to think explicitly about them while making those decisions. Verbal protocols facilitated verbalisation of those decisions *while* they were made and later on provided an important tool for facilitating situated recall. Although initially, all tutors had a clear understanding of and an ability to identify the factors related to subject domain taught, e.g. the difficulty of the material or correctness of student answer, they were much less willing or fluent at diagnosing and talking about factors related to student's affective states. However, after an initial familiarisation period, involving up to two sessions, their willingness to engage in situational analysis and the fluency of their reports increased, while the tentativeness in identifying student behaviours at fine level of details decreased. This was evidenced primarily in the increased speed at which they engaged in the task, the fluency and quality of their verbal protocols and in the post-hoc interviews. Another interesting outcome was the tutors' increased attention to giving praise in their feedback, as well as a more targeted attention to possible relationship between the form of students' responses and their mental states.

The use of verbal protocols during the interactions, each of which was followed by semi-structured interviews, allowed the tutors to formulate hypotheses about the possible meanings of the students' different behaviours in terms of cognitive and affective states and to evaluate those first against the appropriateness of their feedback and then during subsequent tutoring sessions with further students. Finally, post-task walkthroughs were used with the tutors, during which situated recall was facilitated through replay of the video-recorded screens and verbal protocols. The fact that the tutors were given the opportunity to inspect their selection of situational factors and to correct them gave them an opportunity to assess the consistency of their interpretations and further, to analyse those situations where they did not agree with themselves, leading, in some tutors' own words, to deep reflection and grounding of their understanding of (a) what matters to them the most in tutoring situations and (b) the kinds of tutoring they want to be able to deliver *ideally*. The appreciation of the tutors' involvement in the LeAM's KE process was reflected in their request for a tutoring system for tutors, through which they could rehearse and perfect their understanding of the different nuances of educational interactions along with their pedagogical feedback and which they could also use to train novice tutors.

Although the realisation of the potential value of KE methods used to inform an intelligent tutoring system such as LeAM was very inspirational, the methods used, specifically, the way in which they were used, was fundamentally research-centric. The studies were aimed specifically and exclusively to establish some ground truths about very particular kinds of educational interactions for the purpose of creating knowledge representations to underpin the system's learner modelling and natural language dialogue capabilities. As such the tutors participating in the LeAM studies were in essence merely willing informants for and testers of the technological design ideas. Because of the complexity of the studies' set up the tools and the methods used in the study did not lend themselves readily for independent use by the tutors.

The importance of practitioner independence in generating evidence of their practices is emphasised throughout the EBP literature, where it is often accompanied by the rhetoric of *action research* [4] and the call for practitioners as researchers of their own practices. This rhetoric was used to underpin the design of the TARDIS system – a serious game for coaching young people in job interview skills through interactions with intelligent conversational agents able to react to social cues and complex mental states as detected and modelled by TARDIS' user modelling tools [14]. The TARDIS project took LeAM's insights forward, by employing KE methods throughout. Apart from the goal of informing the design of the game, the goal was also to inform the *design of use* of such a game in real contexts of youth employment associations across Europe. Independence of use by practitioners as facilitators of this game was key. In TARDIS, KE was used as the basis for developing practitioners' self-observation and self-reporting skills, which were then built on in the formative evaluation studies, in which the practitioners increasingly participated as researchers, with the support by researchers being gradually removed. The whole process was divided into three stages, roughly corresponding to the three years of the project. The first stage (*familiarisation*) involved gradual preparation and training of practitioners in the application of knowledge elicitation for the purpose of knowledge representation in the domain of job interview training.

Post-hoc walkthroughs, using video replays of practice of job interview sessions between youngsters and practitioners were used to (a) access practitioners' expert knowledge to be represented in TARDIS; (b) allow the practitioners to make overt to themselves, and to the researchers, the types of knowledge and interpretations that are of particular interest in the context of job interview skills coaching and (c) allow the practitioners to reflect on their and the youngsters' needs, leading up to the specification of the necessary and sufficient elements of a technology-enhanced learning environment able to support those needs. This specification was captured in the form of requirements and recommendations, while the reflections were recorded as practitioners' videos annotations in an off-the-shelf tool called *Elan* (https://tla.mpi.nl/tools/tla-tools/elan/).

The second stage (*testing, critique and design of use*) involved a period of continuous cycles of reflection, observation, design and action scaffolded by researchers and guided by the Persistent Collaboration Methodology [5]. This stage was crucial not only to the TARDIS researchers who were able to implement ever more sophisticated prototypes, but it was also fundamental to the practitioners' growing confidence in providing targeted critique of those prototypes, to their increased independence in using TARDIS and in experimenting with its different set-ups. Crucially, the knowledge self-elicitation skills, developed in the first year, along with their rehearsed focus on the type and form of information needed by the researchers to create the various computational models, provided the practitioners with a structure against which to report their observations and reflections to the researchers and a common language for both. One of the key outcomes of this was a growing sense of co-ownership of the tools and knowledge developed which was reflected in the independent curation of TARDIS tools by the practitioners who participated in the project to other practitioners. As such the participating practitioners became *lead-practitioners* in co-designing with their colleagues the use of TARDIS in their everyday practices. This independence was put to the test and further deepened in the third and final stage of the project, where the practitioners engaged in summative evaluation of the system with minimal support from the researchers (*independent use and research*). As well as being able to use the system independently and to explore new ways in which to utilise it within their existing practices, a key outcome was the practitioners' confidently vocal involvement in the development and testing of a schema for annotating data of youngsters engaging in job interviews. This schema was used directly in the analysis of the TARDIS evaluation data, offering the first such tool for examining job interview skills at the low level of detail needed to build user models and artificial agents in this domain [3].

The practitioners' roles and competencies have evidently changed from those of willing informants (the beginning of the project), through advisors and co-designers of the TARDIS system (middle of the project), to lead-practitioners who initiate projects independently (end of the project). At the core of this change was a gradual shift in the practitioners' way of thinking and viewing the world of their practice. Through engaging in KE and its eventual KR in terms of design recommendations and fine-grained specification of the domain and inferences therein (annotation schema), the practitioners' role in applying technology in their practices changed from that of mere consumers to its co-creators and owners. They demonstrated an ability to think about their domain and practices in terms that are by nature both computational (low level knowledge specification) and design (design of the technology's look-and-feel, functionality, as well as pedagogical design[2]). In other words the practitioners have demonstrated an emergent ability to engage in *computational design thinking.*

5 Conclusions

This paper argued a position that the relationship between AIEd and Education can be strengthened through the application of AI as a methodology for supporting educational evidence-based practices. AI offers to educational practitioners specific instruments for generating evidence of their practices that are inspectable and reproducible by the wider educational community. AI methods of knowledge elicitation and representation can enable practitioners to engage

[2] Note that some researchers in Education view teacher as a design science, e.g. [9]

in computational design thinking and this can engender practitioners independence in defining, creating and inspecting their real-world practices at a low-level of representational detail. Investing in educational practitioners using AI as a methodology is not entirely altruistic insofar as the specificity of the evidence thus generated creates an important opportunity for AIEd to tap into situated knowledge of educational practices in a way that supports the implementation of AIEd systems sustainably and over long-term. Such investment carries a promise of creating a dynamically generated knowledge infrastructure thereby reducing the often prohibitive cost of developing AIEd systems and by lending itself more readily to targeted mining and interpretation by the AIEd researchers and developers. Making the AI methods available to practitioners opens the AIEd research to critical, but informed inspection by some of its end-users and it offers a much needed opportunity to re-interrogate its approaches to connecting with existing educational practice, along with its future goals and aspirations more generally.

References

1. Biesta, G.: Why 'what works' won't work: Evidence-based practice and the democratic deficit in educational research. Educational Theory 57(1) (2007)
2. Bundy, A.: What kind of field is artificial intelligence? In: DAI Research Paper No.305. Department of Artificial Intelligence, University of Edinburgh. (1986)
3. Chryssafidou, E., Porayska-Pomsta: Situated job interview coaching: Evaluating the role and efficacy of TARDIS serious game. Journal of Computers in Human Behavior (in prep)
4. Cohen, L., Manion, L.: Research Methods in Education. 2nd edn. Croom-Helm, Dover, NH (1990)
5. Conlon, T., Pain, H.: Persistent collaboration: methodology for applied AIED. Journal of Artificial Intelligence in Education 7, 219–252 (1996)
6. Davis, R., Shrobe, H., Szolovits, P.: What is knowledge representation? AI Magazine 14(1), 17–33 (1993)
7. Dewey, J.: Experience and Nature. Dover Publications Inc. (1998)
8. Hargreaves, A.: Revitalising educational research: lessons from the past and proposals for the future. Cambridge Journal of Education 29(2), 239–249 (1999)
9. Laurillard, D.: Teaching as a Design Science: Building Pedagogical Patterns for Learning and Technology. Routledge (2012)
10. Mark, M.A., Greer, J.: Evaluation methodologies for intelligent tutoring systems. Journal of Artificial Intelligence in Education 4, 129–153 (1993)
11. Porayska-Pomsta, K., Bernardini, S.: Learner modelled environments. In: Handbook of Digital Technology Research (2013)
12. Porayska-Pomsta, K., Mavrikis, M., D'Mello, S., Conati, C.and Baker, R.: Knowledge elicitation methods for affect modelling in education. International Journal of Artificial Intelligence in Education 22, 107–140 (2013)
13. Porayska-Pomsta, K., Mavrikis, M., Pain, H.: Diagnosing and acting on student affect: the tutors perspective. International Journal of Artificial Intelligence in Education 18, 125–173 (2008)
14. Porayska-Pomsta, K., Rizzo, P., Damian, I., Baur, T., Andre, E., Sabouret, N., Jones, H., Anderson, K., Chryssafidou, E.: Who's afraid of job interviews? definitely a question for user modelling. In: Proc. Conference on User Modeling, Adaptation and Personalization (2014)

AIED Is Splitting Up (Into Services) and the Next Generation Will Be All Right

Benjamin D. Nye

Institute for Intelligent Systems, University of Memphis
365 Innovation Dr. Memphis, TN 38152
benjamin.nye@gmail.com

Abstract. Advanced learning technologies are reaching a new phase of their evolution where they are finally entering mainstream educational contexts, with persistent user bases. However, as AIED scales, it will need to follow recent trends in service-oriented and ubiquitous computing: breaking AIED platforms into distinct services that can be composed for different platforms (web, mobile, etc.) and distributed across multiple systems. This will represent a move from learning platforms to an ecosystem of interacting learning tools. Such tools will enable new opportunities for both user-adaptation and experimentation. Traditional macro-adaptation (problem selection) and step-based adaptation (hints and feedback) will be extended by meta-adaptation (adaptive system selection) and micro-adaptation (event-level optimization). The existence of persistent and widely-used systems will also support new paradigms for experimentation in education, allowing researchers to understand interactions and boundary conditions for learning principles. New central research questions for the field will also need to be answered due to these changes in the AIED landscape.

1 Introduction

Initial efforts to bring learning technology into schools faced hardware hurdles, such as insufficient computing resources. Later efforts encountered serious barriers related to matching technology to teachers' beliefs, pedagogy, and resource constraints. While all of these barriers are still relevant, learning technology is endemic in higher education and has made significant footholds in K-12 schools, with estimates of 25-30% of science classes using technology as early as 2012 (Banilower, Smith, Weiss, Malzahn, Campbell, & Weis, 2013). Correspondingly, an influx of investment into educational technology has occurred, with online learning doubling from a $50b industry to a $107b industry in only three years (Monsalve, 2014).

Future barriers will not be about getting learning technology into schools: they will be about competing, integrating, and collaborating with technologies already in schools. This is not an idle speculation, as it is already occurring. In a recent multi-year efficacy study to evaluate a major adaptive learning system, some teachers started using grant-purchased computers to use other math software as well (Craig, Hu, Graesser, Bargagliotti, Sterbinsky, Cheney, & Okwumabua, 2013). After working for many years to get teachers to use technology, the point may come where they are using so many technologies that it is difficult to evaluate an intervention in isolation.

Some research-based artificial intelligence in education (AIED) technologies have already grown significant user bases, with notable examples that include the Cognitive Tutor (Ritter, Anderson, Koedinger, & Corbett, 2007), ALEKS (Falmagne, Albert, Doble, Eppstein, & Hu, 2013), and ASSISTments (Heffernan, Turner, Lourenco, Macasek, Nuzzo-Jones, & Koedinger, 2006). Traditionally non-adaptive systems with large user bases, such as Khan Academy and EdX, have also started to add basic adaptive learning and other intelligent features (Khan Academy, 2015; Siemens, 2013).

Large-scale online platforms are not just the future of learning, but they are also the future of research. Traditional AIED studies have been limited to dozens to hundreds of participants, sometimes just for a single session. While such studies will remain important for isolating new learning principles and collecting rich subject data (e.g., biometrics), large-scale platforms could be used to run continuously-randomized trials across thousands of participants that vary dozens or even hundreds of parameters (Beck and Mostow 2006; Liu, Mandel, Brunskill, & Popovic, 2014). Even for AIED work not based on such platforms, it is increasingly feasible to "plug in" to another system, with certain systems serving as active testbeds for 3rd-party experiments (e.g., ASSISTments and EdX).

The difference is qualitative: rather than being limited to exploring a handful of factors independently, it will be possible to explore the relative importance of different learning principles in different contexts and combinations. In many respects, this means not just a change to the systems, but to the kinds of scientific questions that can and will be studied. These opportunities raise new research problems for the field of AIED. A few areas related areas will reshape educational research: Distributed and Ubiquitous Intelligent Tutoring Systems (ITS), Four-Loop User Adaptation, AI-Controlled Experimental Sampling, and Semantic Messaging. Some new frontiers in each of these areas will be discussed.

2 Distributed and Ubiquitous AIED

As implied by the title, AIED technologies are approaching a juncture where many systems will be splitting up into an ecosystem of reusable infrastructure and platforms. The next generation of services will be composed of these services, which may be hosted across many different servers or institutions. More specifically, we may be reaching the end of the traditional four-component ITS architecture with four modules: Domain, Pedagogy, Student, and Communication (Woolf, 2010). While the functions of all these modules will still be necessary, there is no reason to think that any given ITS must *contain* all these components, in the sense of building them, controlling them, or owning them. The future for ITS may be to blow them up so that each piece can be used as a web-service for many different learning systems.

With respect to other online technologies, learning technology is already behind. On even a basic blog site, a user can often log in using one of five services (e.g., Google, Facebook), view adaptively-selected ads delivered by cloud-based web services that track users across multiple sites, embed media from anywhere on the internet, and meaningfully interact with the site on almost any device (mobiles, tablets, PC). In short, most web applications integrate and interact with many other web services, allowing them to be rapidly designed with robust functionality and data that no single application would be able to develop and maintain.

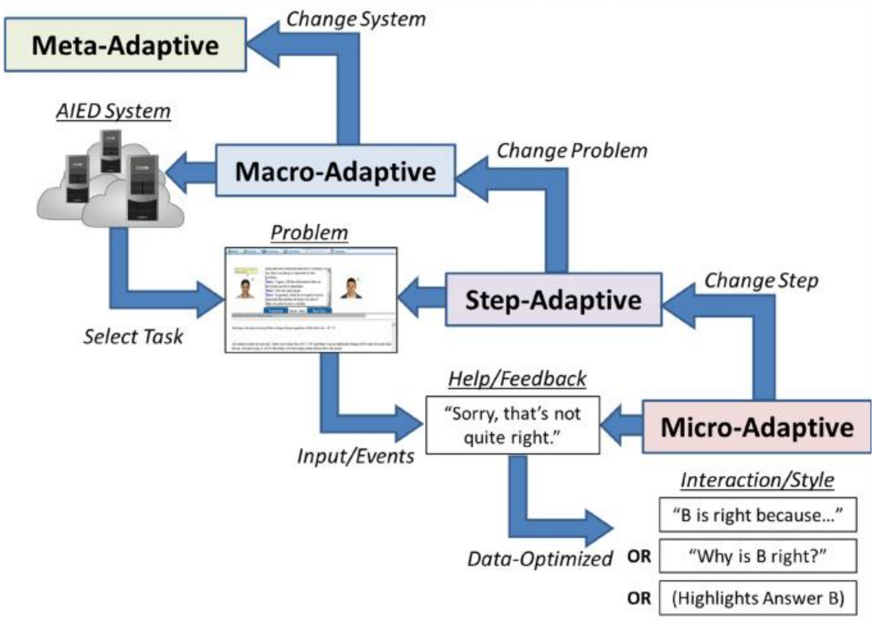
From the standpoint of AIED, moving in this direction is an existential necessity. Without pooling capabilities or sharing components, serious academic research into educational technologies may be boxed out or surpassed by the capabilities of off-the-shelf systems, many of which will have closed architectures. Unfortunately, while industry research can offer powerful results, competing pressures can lead to under-reporting: publishing research is costly, time-consuming, and can risk disclosing trade secrets or unfavorable empirical findings. While some companies make the investment to generalize their research, many others do not. By comparison, academic institutions and research-active commercial systems should be motivated to share and combine technologies to build more effective and widely-used learning technology. This model of collaborative component design stands alone in making platforms that co-exist with major commercial endeavors, such as web-browsers (FireFox), operating systems (Linux), and statistical packages (R; R Core Team, 2013). Moreover, service-oriented computing allows for a mixture of free research development and commercial licensing of the same underlying technologies.

The benefits of moving toward service-oriented AIED will be substantial. First, they should enable AIED research to deeply specialize, while remaining widely applicable due to the ability to plug in to other platforms with large and sustained user bases. In such an ecosystem, user adaptation will be free to expand beyond the canonical inner loop and outer loop model (VanLehn, 2006). Composing and coordinating specialized AIED services will also demand greater standardization and focus on data sharing between systems. While this process may be painful initially, standards for integrating data across multiple systems would enable the development of powerful adaptation, analytics, and reporting functionality that would greatly reduce barriers for developing AIED technology and studying its effects on learners.

3 Four-Loops: Above Outer Loops and Under Inner Loops

One implication of scaling up AIED and moving beyond the standard four-component ITS model is that adaptation to users may become prevalent at grain sizes larger and smaller than traditional ITS. VanLehn (2006) framed the adaption from tutoring systems as consisting of an outer loop (selecting problems) and an inner loop (providing help and feedback on specific problem steps). These are often referred to as "macro-adaptivity" and "step-based adaptivity." However, recent developments have shown the first steps toward "meta-adaptivity," where the system adapts to the user by shifting the learner to an entirely different ITS system (which may then adapt to the user differently). Likewise, research on "micro-adaptivity" has looked at the benefits for using data to fine-tune interactions below the problem step level (e.g., keystroke-level inputs, emotion detection, presentation modes or timing of feedback). This implies a four-loop model for user adaptation, as shown in Figure 1.

Fig. 1. Four-Loop User Adaptation



3.1 Meta-Adaptation: Handoffs Between Systems

Meta-adaptation has only become possible recently, due to increasing use and maturity of AIED technology. In the past, learning technologies such as ITS were trapped in sandboxes with no interaction. Due to service-oriented approaches, systems have taken the first steps toward real-time handoffs of users between systems. For example, in the recent Office of Naval Research STEM Grand Challenge, two out of four teams integrated multiple established adaptive learning systems: Wayang Outpost with ASSISTments (Arroyo, Woolf, & Beal, 2006; Heffernan et al., 2006) and AutoTutor with ALEKS (Nye, Windsor, Pavlik, Olney, Hajeer, Graesser, & Hu, in press). Other integration efforts are also underway as part of the Army Research Lab (ARL) Generalized Intelligent Framework for Tutoring (GIFT) architecture, which is built to integrate external systems (Sottilare, Goldberg, Brawner, & Holden, 2012) and version of AutoTutor has also been integrated with GIFT.

These initial integrations represent the first steps toward meta-adaptation: transferring the learner between different systems based on their needs and performance. This type of adaptation would allow learners to benefit from the complementary strengths of multiple systems. For example, learners that benefit most from animated agents might be sent to systems such agents (i.e., trait-based adaptation). Alternatively, different types of learning impasses or knowledge deficiencies may respond best to learning activities in different systems (i.e., state-based adaptation). One problem that this approach might mitigate is the issue of wheel spinning, where an adaptive system detects that it cannot serve the learner's current needs (Beck & Gong, 2013). Meta-adaptation might also mean referring the learner to a human instructor, tutor, or peer. In general, meta-adaptation would focus on passing students and knowledge between different adaptive learning contexts (both AI-based and human).

Meta-adaptation is the maximum possible grain size, which makes it somewhat different from standard adaptation because users are transferred to an entirely different system. This type of adaptation likely requires either distributed adaptation or brokered adaptation. Distributed adaptation would involve individual systems deciding when to refer a learner to a different system and possibly trusting the other system to transfer the student back when appropriate. This would be analogous to doctors in a hospital, who rely on networks of specialists who share charts and know enough to make an appropriate referral, but may use their own judgment about when and how they make referrals. On the converse, brokered adaptation would require a new type of service whose purpose is to monitor student learning across all systems (i.e., a student model integrator) and make suggestions for appropriate handoffs. This service would be consulted by each participating AIED system, probably as part of their outer loop. In the long term, such a broker may be an important service, because it could help optimize handoffs and ensure that students are transferred appropriately. Such brokers might also play a role for learners to manage their data and privacy settings. Other models for coordinating handoffs might also emerge over time.

3.2 Micro-Adaptation: Data-Optimization and Event Streams

In addition to adaptivity at the largest grain size (selecting systems), research on the smallest grain sizes (micro-adaptation) is also an important future area. Micro-adaptation involves optimizing for and responding to the smallest level of interactions, even those that are not associated with a traditional user input on a problem step. For anything but simple experiments, this type of optimization and adaptation is too fine-grained and labor-intensive to perform by hand at scale, meaning that it will need to rely on data-driven optimizations such as reinforcement learning. Chi, Jordan and VanLehn (2014) used reinforcement learning to optimize dialog-based ITS interactions in the Cordillera system for Physics, which showed potential gains of up to 1σ over poorly-optimized dialog or no dialog. Dragon Box has taken a related approach by optimizing for low-level user interface and click-level data, by applying trace-based models to find efficient paths for learning behavior and associated system responses (Andersen, Gulwani, & Popovic, 2013).

These lines of research represent the tip of the iceberg for opportunities for micro-adaptation. A variety of low-level data streams have not yet been leveraged. Continuous sensor data, such as emotion sensors or speech input waveforms, may present rich opportunities for exploring fine-grained user-adaptation based on algorithmic exploration of possible response patterns. Low-level user interface optimization may also help improve learning, such as human-computer interaction design or keystroke-level events or mouse-over actions (i.e., self-optimizing interfaces).

Both the strength and the drawback of micro-optimization is that it is will tightly fit the specific user interface or content (even down to specific words in text descriptions). Optimizing for a particular presentation of a problem can lead to learning efficiency gains by emphasizing parts that are salient to learning from that specific case, while skipping or downplaying other features. However, micro-level optimization will likely suffer from versioning issues (e.g., changes to small problem elements potentially invalidating prior data and policies) and also transferability issues (e.g., an optimized case not transferring well from a desktop to a mobile context). Solutions to weight the relevance of prior data will be required to address issues related to altered problems or new contexts (e.g., mobile devices, classroom vs. home, different cultural contexts).

4 AI-Controlled Experimental Sampling

Techniques for micro-adaptation may also reshape experimental methods. Artificial intelligence can play a major role in the experimental process itself, which is a type of efficient search problem. Educational data mining research has already started looking at dynamically assigning subjects to different learning conditions based on multi-armed bandit models (Liu, Mandel, Brunskill, & Popovic, 2014). Multi-armed bandit models assume that each treatment condition is like a slot machine with different payout distributions (e.g., student learning gains). These models are common in medical research, where it is important to stop treatments that show harms or a consistent lack of benefit. They allow building intelligent systems that explore new strategies, while pruning ineffective ones.

The field is only taking its first baby steps for these types of experimental designs. Fundamental research is needed to frame and solve efficient-search problems present in AIED experiments. Based on varying different parameters and interactions in the learning experience, learning environments can search for interpretable models that predict learning gains. In the long term, models for automated experimentation may even allow comparing the effectiveness of different services or content modules, by randomly selecting them from open repositories of content.

The most difficult aspect of this problem is likely to be the interpretability. While multi-arm bandit models can be calibrated to offer clear statistical significance levels between conditions, models that traverse the pedagogical strategy space are often too granular to allow for much generalization. For example, some popular models for large learning environment focus on efficient paths or traces of learning behavior and associated system responses (Andersen, Gulwani, & Popovic, 2013). Unfortunately, these models are often not easily generalizable: they may capture issues tied to the specific system or may tailor instruction to specific problems so tightly that it is difficult to infer theoretical implications (Chi, Jordan & VanLehn, 2014).

New techniques are needed that can automatically explore the space of pedagogical designs, but that can also output interpretable statistics that are grounded in theories and concepts that can be compared across systems. This is a serious challenge that probably lacks a general algorithmic solution. Instead, such mappings will probably be determined by the constraints of learning and educational processes. A second major challenge is the issue of integrating expert knowledge with statistically-sampled information. Commonly, expert knowledge is used to initially design a system (e.g., human-defined knowledge prerequisites), which is later replaced by a statistically-inferred model after enough data is collected. However, in an ideal world, these types of heterogeneous data would be gracefully integrated (e.g., treating expert knowledge as Bayesian prior weights). Future research in AIED will need to identify where this sort of expert/statistical hybrid modeling is needed, and match these problems with techniques from fields of AI and data modeling that specialize in these issues. Ultimately, a goal of this work should be to blur the lines between theory and practice by building systems that can both report and consume theoretically-relevant findings.

5 Semantic Messaging: Sharing Components and Data

To share technology effectively, AIED must move toward open standards for sharing data both after-the-fact (i.e., repositories) and also in real-time (i.e., plug-in architectures). The first steps in these directions have already been taken. Two notable data repository projects with strong AIED roots exist: the Pittsburg Science for Learning Center (PSLC) DataShop (Koedinger, Baker, Cunningham, Skogsholm, Leber, & Stamper, 2010) and the Advanced Distributed Learning (ADL) xAPI standards for messaging and learning record stores (Murray & Silvers, 2013). The IMS Global Specifications are also a move in this direction (IMS Global, 2015).

Due to solid protocols in messaging technologies, the technical process of exchanging data between systems at runtime is not onerous. The larger issue is for a receiving system to actually apply that data usefully (e.g., understand what it means). Hidden beneath this issue is a complex ontology alignment problem. In short, each learning technology frames its experiences differently. When these experiences and events are sent off to some other system, the designers of each system need to agree about what different semantics mean. For example, one system may say a student has “Completed” an exercise if they viewed it. Another might only mark it as “Completed” if the learner achieved a passing grade on it. These have very different practical implications. Likewise, the subparts of a complex activity may be segmented differently (e.g., different theories about the number of academically-relevant emotions). While efforts have been made to work toward standards, this seldom solves the problem: the issue with standards is that there tends to be so many of them.

So then, ontology development must play a key role for the future of ITS interoperability. There are multiple ways that this might occur. Assuming the number of standards is countable, it would be sufficient to have an occasional up-front investment to develop and update explicit mappings between ontologies by hand. While this is low-tech, it works when the number of terms is fairly small. For larger ontologies of AIED behavior and events, it may be possible to align ontologies by applying both coding systems to a shared task (e.g., build benchmark tasks that are then marked up with messages derived from that ontology).

By collecting data on messages from benchmark tasks, it may be possible to automate much of the alignment between ontologies, particularly for key aspects such as assessment. Research on Semantic Web technologies is also very active, and may offer other effective solutions to issues of ontology matching and alignment (Shvaiko & Euzenat, 2013). The final approach is to simply live without standards and allow the growth of a folksonomy: common terms that are frequently used. These terms can then become suggested labels, with tools that make their use more convenient and prevalent. The one approach that should *not* be taken is to try to develop a super-ontology or new top-down standard for the types of information that learning systems communicate. While there are roles for such ontologies, top-down ontologies have never achieved much support within research or software development communities.

6 Closing Remarks

The future for AIED should be a bright one: expansion of learning software into schools will ultimately result in unprecedented diversity and size of user bases. The areas noted in this paper are only the first wave for new AIED opportunities. In time, it will be possible to explore entirely new classes of questions, such as mapping out continuous, multivariate functional relationships between student factors and pedagogical effectiveness of certain behaviors. Systems such as personal learning lockers for data would allow for longitudinal study of learning over time, either in real-time or retrospectively. A major game-changer for future learning research will probably be data ownership and privacy issues: data will exist, but researchers will need to foster best-practices for data sharing, protection, and archiving.

With this wealth of data, researchers will be able to connect learning to other relationships and patterns from less traditional data sources. In 20 years, the range of commonly-available sensor data will be dizzying: geolocation, haptic/acceleration, camera, microphone, thermal imaging, social ties, and even Internet-of-Things devices such as smart thermostats or refrigerators. Moreover, the ecosystem of applications leveraging this data will likewise be more mature: your phone might be able to tell a student not only that their parents left them a voicemail, but that they sounded angry. This event might then be correlated with a recent report card, and the consequences of the interaction might be analyzed. Learning is a central facet of the human experience, cutting across nearly every part of life. To that end, as life-long learning becomes the norm, the relationship between life and learning will become increasingly important. By consuming and being consumed in a distributed and service-oriented world, AIED will be able to play a major role in shaping both education and society.

References

1. Andersen, E., Gulwani, S., & Popovic, Z. (2013). A trace-based framework for analyzing and synthesizing educational progressions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 773-782). ACM.

2. Arroyo, I., Woolf, B. P., & Beal, C. R. (2006). Addressing cognitive differences and gender during problem solving. *International Journal of Technology, Instruction, Cognition and Learning*, *4*, 31-63.
3. Banilower, E . R., Smith, P. S., Weiss, I. R., Malzahn, K. A., Campbell, K. M., & Weis, A. M. (2013). *Report of the 2012 national survey of science and mathematics education*. Chapel Hill, NC: Horizon Research, Inc.
4. Beck, J. E., & Gong, Y. (2013, January). Wheel-spinning: Students who fail to master a skill. In *Artificial Intelligence in Education* (pp. 431-440). Springer Berlin.
5. Chi, M., Jordan, P., & VanLehn, K. (2014). When Is Tutorial Dialogue More Effective Than Step-Based Tutoring? In *Intelligent Tutoring Systems* (pp. 210-219). Springer: Berlin.
6. Craig, S. D., Hu, X., Graesser, A. C., Bargagliotti, A. E., Sterbinsky, A., Cheney, K. R., & Okwumabua, T. (2013). The impact of a technology-based mathematics after-school program using ALEKS on student's knowledge and behaviors. *Computers & Education*, *68*, 495-504.
7. Falmagne, J. C., Albert, D., Doble, C., Eppstein, D., & Hu, X. (2013). *Knowledge Spaces: Applications in Education*. Springer Science & Business Media.
8. Graesser, A. C. (2011). AutoTutor. In P. M. McCarthy, & C. Boonthum (Eds.), *Applied natural language processing and content analysis: Identification, investigation and resolution*. Hershey, PA: IGI Global.
9. Heffernan, N. T., Turner, T. E., Lourenco, A. L., Macasek, M. A., Nuzzo-Jones, G., & Koedinger, K. R. (2006). The ASSISTment Builder: Towards an Analysis of Cost Effectiveness of ITS Creation. In *FLAIRS Conference* (pp. 515-520).
10. IMS Global (2015). Learning Tools Interoperability. Retrieved from: http://www.imsglobal.org/lti/index.html on Feb 22, 2015.
11. Khan Academy (2015). About Khan Academy. Retrieved www.khanacademy.org/about on Feb. 22, 2015.
12. Koedinger, K. R., Baker, R. S., Cunningham, K., Skogsholm, A., Leber, B., & Stamper, J. (2010). A data repository for the EDM community: The PSLC DataShop. In C. Romero, S. Ventura, M. Pechenizkiy, & R. S.J.d. Baker (Eds.) *Handbook of educational data mining*, 43-53.
13. Liu, Y. E., Mandel, T., Brunskill, E., & Popovic, Z. Trading Off Scientific Knowledge and User Learning with Multi-Armed Bandits. *Educational Data Mining (EDM) 2014*. 161-168. Springer
14. Monsalve, S. (2014). A venture capitalist’s top 5 predictions for 2015. *Fortune*. Retrieved http://fortune.com/2014/12/29/a-venture-capitalists-5-predictions-for-2015/ on Feb. 22, 2015.
15. Mostow, J., & Beck, J. (2006). Some useful tactics to modify, map and mine data from intelligent tutors. *Natural Language Engineering*, *12*(02), 195-208.
16. Murray, K., & Silvers, A. (2013). A learning experience. *Journal of Advanced Distributed Learning Technology*, *1*(3-4), 1-7.
17. Murray, T., Blessing, S., & Ainsworth, S. (2003). *Authoring tools for advanced technology learning environments: Toward cost-effective adaptive, interactive and intelligent educational software*. Springer.
18. Nye, B. D., Windsor, A., Pavlik, P. I., Olney, A., Hajeer, M., Graesser, A. C., & Hu, X. (In Press). Evaluating the Effectiveness of Integrating Natural Language Tutoring into an Existing Adaptive Learning System. In *Artificial Intelligence in Education (AIED) 2015*.
19. R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0

20. Ritter, S., Anderson, J. R., Koedinger, K. R., & Corbett, A. (2007). Cognitive Tutor: Applied research in mathematics education. *Psychonomic bulletin & review*, *14*(2), 249-255.
21. Shvaiko, P., & Euzenat, J. (2013). Ontology matching: state of the art and future challenges. *Knowledge and Data Engineering, IEEE Transactions on*, *25*(1), 158-176.
22. Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, 0002764213498851.
23. Sottilare, R. A., Goldberg, B. S., Brawner, K. W., & Holden, H. K. (2012). A modular framework to support the authoring and assessment of adaptive computer-based tutoring systems (CBTS). In *Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) 2012*.
24. VanLehn, K. (2006). The behavior of tutoring systems. *International Journal of Artificial Intelligence in Education*, *16*(3), 227-265.
25. Woolf, B. P. (2010). Building intelligent interactive tutors: Student-centered strategies for revolutionizing e-learning. Morgan Kaufmann.

Education still needs Artificial Intelligence to support Personalized Motor Skill Learning: Aikido as a case study

Olga C. Santos

aDeNu Research Group. Artificial Intelligence Dept. Computer Science School, UNED. Calle Juan del Rosal, 16. Madrid 28040. Spain
http://adenu.ia.uned.es
ocsantos@dia.uned.es

Abstract. Motor skill learning is hardly considered in current AIED literature. However, there are many learning tasks that require consolidating motor tasks into memory through repetition towards accurate movements, such as learning to write, to draw, to play a musical instrument, to practice a sport technique, to dance, to use sign language or to train for surgery. The field of Artificial Intelligence (AI) needs new sap to cope with the challenges in the Educational (ED) domain aimed to support psychomotor learning. This new sap can be provided by novel interactive technologies around the Internet of the Things that deal with Quantified-self wearable devices, 3D modelling, Big Data processing, etc. The paper aims to identify opportunities and challenges for AI + ED that can be discussed during the workshop. Some of the issues raised are illustrated within a case study instantiated in the Aikido practice, a defensive martial art that involves learning skilled movements by training both the body and the mind, and which is not only part of extra-curricular activity in many schools, but has also been reported of value for teaching in STEM (Science, Technology, Engineering and Mathematics) education, in particular, some laws of mechanics.

Keywords: motor skill learning, psychomotor domain, artificial intelligence, education, Internet of the Things, personalization, Aikido, STEM.

1 Introduction

Motor skill learning can be defined as achieving the ability to perform a function acquired with practice that requires body and/or limb movement to accomplish the goal of an action or task [1]. Although it is not a new concept [2], up to my knowledge (grounded by a review of the papers published in the International Journal of Artificial Intelligence in Education (IJAIED) and which is reported elsewhere [3]), the physical aspects of learning have been hardly considered in the AIED research. Nevertheless, consolidating specific motor tasks into memory through repetition (thus, creating long-term muscle memory for a given task) is very relevant in diverse educational scenarios that support learning processes involving not only brain activity, but also physical activity, such as learning to write, to draw, to play a musical instrument, to practice a sport technique, to dance, to use sign language or to train for surgery that require long-term physical training, as reported in [3]. In these situations, learners have to train by repeating basic and very specific movements till they learn the best way to carry them out effectively without conscious effort. It has to be remarked here that learning physical skills (i.e., the proficiency of individual movements, also called sensomotor habits [4]) goes beyond mere muscle memory, but involve blending motor skills and cognitive, meta-cognitive and affective skills. In fact, psychomotor skills cannot be acquired by multiple repetitions of given motor pattern without considering the importance of feedback between cognitive processes and motor actions [5]. However, the focus of the discussion that this paper aims to bring to the workshop is mainly on how the physical part related to the psychomotor learning domain (which deals with physical movement, coordination and the use of the motor skill areas [6]) can be supported from an AIED perspective, both in 1) the modelling of the learner physical interaction, and 2) the provision of the required personalized support during the learning. In my view, this is a new dimension that is worth to be explored by combining AI + ED research. The cognitive, meta-cognitive and affective dimensions are already being widely addressed in AIED literature.

In addition, at this point in time, technology has evolved in such a way that it can monitor the movements carried out by the learners through diverse types of sensors (e.g., inertial, optical, position, electromyography, etc.) and timely feedback can be provided through diverse actuators (such as resistance, force, vibration, etc. as well as servo motors) to help the learner improve the performance of the corresponding movement. Quantified-self approaches (based on data gathered from wearable devices such as electronic bracelets and intelligent t-shirts) allow personal awareness and reflection for behavioral monitoring in many situations, such as physical exercise or affective support. Big Data allows processing real time data streams gathered from heterogeneous information sources. 3D models of real objects can be produced with low-cost scanners and printers. These technologies (among others) support the so called Internet of the Things (IoT), that is, the connection of *physical things* to the Internet, which makes possible to access remote sensor data and to control the physical world from a distance [7]. In this context, the do-it-yourself movement supports non-experts in getting familiar with these novel interaction technologies and in being able to build ad-hoc electronic components for their own needs. Thus, AIED researchers can take advantage of this supportive context so the learning curve of integrating above technologies with AI techniques can be feasible for the field.

As a result, this paper proposes to explicitly open a new research line for the AIED field where ED can benefit from AI techniques enriched with emerging novel interactive technologies around the Internet of the Things. This new research direction, framed within the psychomotor learning domain, requires a shift towards supporting physical practice (i.e., training) rather than supporting instructional teaching. This implies that the physical actions carried out while practicing need to be monitored, modelled and, when needed, corrected, to achieve successful motor skill learning (i.e., skills learning at a physical level).

2 Opportunities and Challenges

As discussed in [3], the synergy of Artificial Intelligence techniques with novel interactive technologies opens new opportunities for researching the physical (i.e., corporal) aspects of learning. For instance, it seems to be possible to provide intelligent real time feedback to scaffold physical skill learning by using sensors, actuators, 3D scanning and modelling, data streams processing, etc. And in order to improve performance, tangible scaffolding could be provided to guide motor skill learning in a personalized way through embodiment technology. A case study that illustrates some issues involved is outlined in Section 3.

In any case, by integrating novel interactive technology, the foreseen goal is that AIED researchers can produce systems that sense the learner's corporal behavior as she learns specific skilled movements, and then guide the learner on how to react in an optimal way (taking into account the learner's current performance, corporal features and the particularities of the specific movement to perform) by providing personalized feedback during the learning process (rather than just giving directions of what to do and how to do, as in traditional AIED intervention approaches). Procedural learning in terms of motor skill is usually difficult to explain by the instructor and to understand by the learner. In fact, this procedural tutoring support is of major relevance in the case of novice learners, as they might get into a wrong habit if no timely feedback is provided to them while practicing by their own and, thus, they cannot understand why the movement is not correct.

In order to build procedural learning systems that can personalize motor skill learning, both AI and ED research need to revise the application of their theoretical and methodological approaches to the particularities of the psychomotor learning domain. From the AI point of view, there is a need for modelling the individual functional and corporal features, her interaction and the accurate movement, by processing the simultaneously and continuously data streams produced by diverse and heterogeneous sensors, and then controlling the robotics to physically deliver the intervention to the learner. From the ED point of view, the focus has to be put on identifying what is the most appropriate intervention in each case (considering cognitive, meta-cognitive, affective and behavioral dimensions) and when and how it should be delivered in order to make a positive impact in the learning process.

Therefore, as discussed in [3], there exist challenges regarding 1) modelling and representing the movements of the learner by building the learner physical interaction model as well as the accurate movement model, and 2) providing the appropriate personalized physical support in the most efficient way for each learner in each training context. More specifically, regarding the modelling of movements, there seem to be challenges related to: i) detecting the physical interaction, ii) modelling the movements to be trained, iii) error diagnosing and intervention modelling, and iv) modelling the learner. In turn, regarding the provision of the appropriate personalized physical support, challenges might exist in order to: i) deciding upon adaptation, ii) evaluating the user activity, iii) visualization of movement performance, and iv) sharing progress and social learning.

3 A case study for AI + ED: supporting personalized psychomotor learning in Aikido

In order to facilitate the discussion on existing challenges for AI + ED to support personalized motor learning skill learning, a case study is presented. This case study focuses on Aikido martial art. Since it might surprise the reader the selection of this domain from an ED perspective, first some of the reasons for its selection are discussed. Then, some technological advances that can help AI to provide personalized motor skill training within the Aikido psychomotor learning domain are presented. They intend to include in the AIED research agenda ideas that can be explored.

3.1 Aikido & ED: more than just a psychomotor learning domain

Aikido is a non-aggressive Japanese martial art that consists of entering and turning movements that redirect the momentum of an opponent's attack, and a throw or joint lock that terminates the technique [8]. The word is formed by Ai (coordination, accord, harmony, blending), Ki (psychological energy, spirit, universal force) and Do (way of life, philosophy of living) [9]. It is guided by defending oneself while also protecting the attacker from injury. In fact, it is based on the principle that in order to control an attacker, the defender must meet the attack in a state of perfect balance [10]. Properly carrying out the technique requires years of training by repeating over and over the sequence of movements that makes up each Aikido technique.

Martial arts do not only involve complex manipulations of human anatomy and physiology [9], but they aim to train both the body and the mind, since training consist of improving mental disposition and motor skills (i.e., fitness and coordination) [4]. According to these authors [4], the technique of self-defense can be defined as a specific sequence of movements constituting a partial or total resolving of various dynamic situations. These movements imply eccentric and concentric muscle work, rotation of the trunk and hips, translocation of the body mass center and adequate leg work. Interplay of muscle tension and relaxation combined with accurate decisions is needed. This requires the development of skills in body movement control that combine mental balance and appropriate motor actions, where the general motor fitness is adjusted to the individual level of motor abilities (i.e., quality is more important than strength). Automation of movements occurs when mental processes are free of controlling individual movements. An ability of psycho-physical self-controls is also required to allow for efficient performance under stressful situations.

Since Aikido practice involves the execution of paired movements between the attacker (*uke*: receiver of the technique) and the defendant (*tory*: doer of the technique), it helps understanding cooperation and timing in movement [11]. Recent studies using electroencephalography and electromyography techniques have shown that the postural control training using Aikido improves psychomotor performance [10].

Nonetheless, the benefits of Aikido go beyond physical fitness and motor abilities. For instance, some studies suggest that Aikido training increases mindfulness [11]. In particular, since practitioners are taught to be mindful of the technique, breathing, balance, center of gravity and their connection to the other person, it may facilitate increasing one's awareness of body position, of others around, practitioner's emotional states and how other people's emotions may affect the Aikido practitioner's emotional states. As compiled by these authors, benefits of increased mindfulness may include better concentration, stronger awareness, improved immune system functioning and decreases in stress related physical symptoms [12, 13]. In this way, Aikido training may enhance awareness and resolution of problematical situations, as during training sessions, the practitioner learns to deal with multiple stressors concurrently, and this is learnt to do in an effective manner while remaining calm, which suggests that Aikido seem to teach practical problem solving and acceptance of circumstances [11]. In this sense, Aikido is one of the more spiritual martial arts as it studies the energy within oneself, her partner and the world through the physical principles of entering, turning and securing, and thus, focuses directly on the energy involved in dealing with one's emotions, perceptions of trust and fear, and conceptions of reality as well as the energy and demands in relating with another human being [14]. In this authors' viewpoint, Aikido can contribute to relationship encounters, conflict resolution, motivation and personal energy by an effective management of energy, improving interpersonal relationships and facilitating stress reduction. Following these ideas, studies have shown that including martial arts such as Aikido in school programs can enhance student's awareness of violence prevention and allow them to react calmly and without panic, reducing violence in schools [15].

In addition to above benefits, Aikido has also potential to be used in education, not only for physical education (i.e., development of motor abilities, mental and physical health benefits, violence reduction…) but also in STEM education (i.e., Science, Technology, Engineering and Mathematics). In this sense, there are studies where some laws of Physics are taught with Aikido practice (see [15]) that show statistically significant improvements in the scores on biomechanics (i.e., mechanics principles of human movement) tests as well as statistically significant correlations between the results in those tests and the performance of the Aikido techniques. From these works, it seems that solid-state mechanics concepts such as the law of momentum conservation, second law of motion for angular motion, centrifugal force and composition of resultant forces and moments of force, can be explained more effectively with the practice of Aikido, facilitating the understanding of how forces act on a person while in translator or rotary motion.

Since the practice of Aikido seems to improve not only motor skills, but also some cognitive abilities (i.e., acquiring the knowledge of mechanics required by the scholar curriculum), this martial art has been chosen to discuss how a psychomotor learning domain like this could benefit from an AIED procedural learning environment. In this sense, some ideas on how to provide some tangible scaffolding when needed to guide motor skill learning in a personalized way using novel interactive technology from the IoT are discussed next. The research question behind is: *How to design and implement a personalized procedural learning environment that can physically train and guide the particular way each learners' body and limbs should move in order to achieve a specific learning goal that is related to improving learners' motor skills acquisition, such as the needs identified in the Aikido practice?*

3.2 Improving AI based personalized motor skill learning in Aikido with novel interactive technologies

The goal of Aikido is to hold the *uke* (attacker) in a compromised and secured position with a minimal amount of effort [17]. To achieve this, Aikido practice involves the manipulation of various joints of the body and is based on effective anatomical principles to subdue a training partner by twisting the limbs or locking up the skeletal system. In order to better understand the body's responses and improve the proficiency of applying specific techniques, anatomical studies on cadavers that investigated the nerves, bones, muscles, tendons and tissues manipulated by each technique have been carried out in the past [9]. However, novel interactive technologies, such as those provided by quantified-self wearable devices, can be used to gather dynamic indicators while making the movement. This can help to understand how the movements are performed and improve training. For instance, the movements carried out by a person can be monitored using diverse types of sensors (inertial, optical, position, physiological, etc.) [18] for real time motion study outside the laboratory [19]. This technology is becoming less and less intrusive, to the point that sensors that allow complex movement patterns tracking are getting embedded directly into clothes [20]. The interaction data streams continuously collected by these sensors in real time need to be processed. Due to its volume, variability and speed, Big Data mining techniques need probably to be applied [21].

In addition, as introduced above, Aikido requires long-term physical training to learn how to carry out the movements in the most efficient way. Very often, the execution of the corresponding techniques involves practitioners moving along a curve and lowering one's center of gravity in order to employ the centrifugal force acting on the opponent and one's own gravity [16]. Forces applied are notably subtle and intricate, and thus, difficult to learn without the direct tutelage by an experienced *sensei* (teacher) [17]. This is not easy to put into practice without being repeatedly told what is done wrong and what should be done right. In order to be able to compare how the movement is performed, a model of the accurate movement needs to be built. In the field of virtual reality, there are works that build virtual skeletal models for videogames from the information collected using wearable technology (e.g., biomechanical or inertial sensors), which both map the movement as well as recognize gestures with AI techniques [22]. The movement controlled by sensors can also be represented in 3D models of the human body [23].

The next step is to provide some guided feedback. Since the situations where the applied techniques are never the same (e.g., the degree and direction of force is different, the position of the *tory* is not always the same, body shape and muscular structure differ from *uke* to *uke*, perception and timing change) the application of the technique must change accordingly [24]. This means that the provided feedback should be personalized to the current situation, including *uke* and *tory* body built. With respect to defining the appropriate feedback to give, an initial proposal can be to provide some tangible scaffolding through embodiment technology that corrects the learner's movements by physically controlling and guiding the movement of the learner till her ideal movements (considering the learner's own body built) are achieved. Feedback with different levels of complexity (simple verification, try again and elaborated) provided through different channels (visual, audio and haptic) [17] should be considered. For instance, in order to provide motor intervention, some works use electromyography sensors (i.e., the measure of the electrical activity produced by the skeletal muscles) to detect movement intentions and help to carry them out through exoskeletons (i.e., physical shells) moved with servo motors [25]. Resistive sensors have also been used to move body parts through vibrations [26]. Inertial sensors and vibro-tactil feedback is also used to replicate referred postures and correct those that are not alike [27]. A forced feedback system to guide fingers movement to improve motor skills when playing the piano has been implemented with a simple exoskeletal robotics [28]. The technology for 3D modelling can be used to build physical prototypes of tangible objects. As an example, combining available technologies, a 3D printed hand has been controlled with Arduino using servomotors [29].

However, guiding the learner by delivering forced haptic feedback when the movement performed does not reflect the reference movement might not be the most appropriate psycho-educational approach to achieve long-term learning, although it might help to increase motivation by contributing to short-term performance [30]. Therefore, there is a need to research the appropriate personalized support to provide. Here, the application of TORMES methodology [31] (or an extended version of it that addresses the particularities required by the psychomotor learning domain and the requirements to sense the environment and provide tangible support) can be of value to model the personalized dynamic psychomotor support to be provided in specific situations. In particular, TORMES extends the design cycle of interactive systems as defined by ISO 9241-210 with the life cycle of e-learning and the layered evaluation of adaptive systems, and combines user centered design methods (which can be applied to gather tacit knowledge from psychomotor experts as well as experienced Aikido teachers and practitioners) with (big) data mining techniques (that can be used to analysis performance indicators regarding the movements carried out gathered from Aikido training sessions, for instance, using wearable devices).

There is a commercial software (i.e., Aikido 3D[1]) that recreates with animated characters the movements of a high degree Aikido black belt using motion capture technology. The goal of this tool is to facilitate visualizing how the techniques are to be carried out, so the learner can see it from different perspectives, in slow motion, zoomed, etc. It provides a technological improvement on top of what takes place in Aikido *dojos* (i.e., training places) around the world, but the approach behind is similar: learner watches how an expert (in this case, an animated character whose behavior has been modelled with the movements of an expert) carries out the technique and then tries to reproduce (imitate) the same movements with a partner. However, an AIED support through a procedural learning environment could improve the learning experience by physically controlling and guiding the movements of the learner when appropriate, so she can correct them till she masters the movements for the technique (considering the learner's own body built and skills, as well as the context where the movement is carried out, including the opponent features). This requires the follow-

[1] https://www.aikido3d.com

ing: 1) sensing the learner's movement and the context in which this movement takes place (e.g., the physical features and abilities of the opponent), 2) comparing it against the accurate movement (e.g., how an expert in the technique would carry the movement out considering the same physical features and abilities of the learner and the opponent), 3) deciding whether it is appropriate to provide the tangible support at this moment (dealing with focusing on short term performance vs. long-term learning), and 4) if appropriate, then provide the tangible support in an effective non-intrusive way, for instance with vibro-tactil feedback through actuators sew on the *Aikidogi* (i.e., the Aikido training uniform).

4 Conclusion

There is a challenge and opportunity to take advantage of AI and ED research to develop personalized procedural systems that can support learners while acquiring psychomotor abilities. Learning and improving motor skills is of relevance in many domains, such as learning to write, to draw, to play a musical instrument, to practice a sport technique, to dance, to use sign language or to train for surgery.

In this paper, the relevance of Aikido practice and the support it can obtain from AIED based procedural learning environments has been discussed for the first time in the literature. In addition, the application of novel interaction technologies that are being used by the Internet of the Things (such as quantified-self wearable devices, big data processing and 3D modelling) to build an AIED procedural learning environment has been proposed by reporting works that partially address some of the technological issues discussed. Although the assimilation of new technologies is always costly, the do-it-yourself movement, which encourages people in creating Internet of the Things applications by their own [32], can simply their learning curve and thus, their usage should be feasible for the AIED research community, provided that many people around the world are taking advantage of them without a wide specialized technological background. In turn, non-specialized users benefit from the feeling of belonging to a community that characterizes this kind of developing culture (as well as the open source and open hardware philosophy underneath it) and receive on-line peer support both on search (i.e., looking for information with the help of web search engines or within specialized repositories) and on demand (i.e., asking in specialized forums).

In addition, it can also be noted here that most of the approaches referenced in the previous section can be controlled by an Arduino based infrastructure. Arduino is an open source electronics prototyping platform, which is based on easy to use hardware and software [33]. As reported in previous work, Arduino can be used to gather contextual information from sensors [34] and deliver ambient intelligent feedback [35].

In summary, the motivation of this paper is to propose a new research direction to the AIED field, where novel interactive technologies enrich Artificial Intelligence techniques to deal with some challenges within the Educational domain. This proposal will be discussed further during the workshop "Les Contes du Mariage: Should AI stay married to ED? A workshop examining the current and future identity of the AIED field" taking place during the 17th International Conference on Artificial Intelligence in Education (AIED 2015). Outcomes from the discussion in the workshop will be included in a paper for the IJAED Special Issue “The next 25 Years: How advanced, interactive educational technologies will change the world” [3].

Acknowledgements

This work has been carried out in the context of the project *“Towards tangible recommender systems: personalizing kinematic recommendations modelling with low cost technology”* (***tangibREC***). Funding has been requested to the Spanish Ministry of Economy and Competitiveness (MINECO) in the Call 2014 for Explore Science and Explore Technology projects.

References

1. Christensson, J. Individualising Surgical Training in a Virtual Haptic Environment. MSc Thesis in Interaction Design. IT University of Göteborg (Sweden), 2005.
2. Fitts, P.M. Perceptual-Motor skills learning. In: Melton AW, eds. Categories of Human Learning. New York, NY: Academic Press Inc., 243-286, 1964.
3. Santos, O.C. Training the Body: the Potential of AIED to support Personalized Motor Skill Learning. Special Issue “The next 25 Years: How advanced, interactive educational technologies will change the world”. International Journal of Artificial Intelligence in Education, 2015 (in preparation).
4. Harasymowicz, J., and Kalina, R.M. Training of psychomotor adaptation – a key factor in teaching self-defence. Archives of Budo, 1, 19-26, 2005Schmidt, R.A. Motor learning and performance. Human Kinetics, Champaign, IL. 1991.
5. Harrow, A. A Taxonomy of Psychomotor Domain: A Guide for Developing Behavioral Objectives; David McKay: New York, NY, USA, 1972.
6. Kopetz, H. Internet of Things, Real-Time Systems Series, 307-323, 2011.
7. Ueshiba, K. The Art of Aikido: Principles and Essential Techniques. Kodansha International, 2004.
8. Seitz, F.C., Olson, G.D., Stenzel, T.E. A martial arts exploration of elbow anatomy: Ikkyo (Aikido's first teaching). Perceptual and Motor Skills, 73, 1227-34 (1991).
9. Bazanova, O.M., Kholodina, N, Kurose-Payet, Y., Payet, J., Nikolenko, E.D., Podoinikov, A. Postural control training using Aikido improves psychomotor performance. International Journal of Psychophysiology, 94(2), 165-165, 2014.Lothtes, J., Hakan, R. and Kassab, K. Aikido experience and its relation to mindfulness: a two-part study. Perceptual and Motor Skills, 116(1), 30-9, 2013.
10. Shapiro, S., Carlson, L., Astin, J., and Freedman, B. Mechanisms of Mindfulness. Journal of Clinical Psychology. 62 (3), 373-386, 2006.
11. Kabat-Zinn, J. Wherever you go there you are: mindfulness meditation in everyday life. New York: Hyperion, 1994.
12. Seitz, F.C., Olson, G.D., Locke, B. and Quam, R. The martial arts and mental health: the challenge of managing energy. Perceptual and Motor Skills, 70(2), 459-464, 1990.
13. Lu, C. Martial Arts, Violence, and Public Schools, Brock Education, Vol 18, 2008, 1-12.
14. Mroczkowski A. Using the Knowledge of Biomechanics in Teaching Aikido. In Injury and Skeletal Biomechanics. Goswami, T. (Ed.). InTech, 2012.

15. Olson, G.D., Cook IV, M., Brooks, L. Aikido's Arm-Lock (Ude-Gatame) Technique: What Tissues are Affected? Journal of Asian Martial Arts, 8(2), 42-49, 1999.
16. Schneider, J., Börner, D., van Rosmalen, P., Specht, M. Augmenting the Senses: A Review on Sensor-Based Learning Support. Sensors, 15, 4097-4133, 2015.
17. Fong, D.T.-P.; Chan, Y.-Y. The Use of Wearable Inertial Motion Sensors in Human Lower Limb Biomechanics Studies: A Systematic Review. Sensors, 10, 11556-11565, 2010.
18. Fleury, A.; Sugar, M.; Chau, T. E-textiles in Clinical Rehabilitation: A Scoping Review. Electronics, 4, 173-203, 2015.
19. Fan, W., Bifet, A. Mining big data: current status, and forecast to the future. SIGKDD Explor. Newsl. 14, 2 (April 2013), 1-5, 2013,
20. Arsenault, D. A Quaternion-Based Motion Tracking and Gesture Recognition System Using Wireless Inertial Sensors. Master of Applied Science in Human-Computer Interaction. Carleton University, 2014.
21. Bae, J., Haninger, K., Wai, D., Garcia, X., Tomizuka, M. A Network-Based Monitoring System for Rehabilitation. The 2012 IEEE/ASME International Conference on Advanced Intelligent Mechatronics, 232-237, 2012.
22. Saotome, M. Aikido and the harmony of nature. Boulogne, France: Serirep. 1986.
23. Gopura, R. A. R. C., Kiguchi, K. EMG-Based Control of an Exoskeleton Robot for Human Forearm and Wrist Motion Assist. IEEE International Conference on Robotics and Automation, 731-736, 2008.
24. Rush, R.P Sensation Augmentation to Relieve Pressure Sore Formation in Wheelchair Users. Eleventh International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS'09), 275-276, 2009.
25. Ding, Z.Q., Luo, Z.Q., Causo, A., Chen, I.M., Yue, K.X. , Yeo, S.H., Ling, K.V. Inertia sensor-based guidance system for upperlimb posture correction. Medical Engineering & Physics, vol. 35 (2), 269-276, 2013.
26. Datta, S. Forced fingers, Available from Github: https://github.com/dattasaurabh82/Forced-Fingers
27. Huluta, E., da Silva, R.F., de Oliveira, T.E.A. Neural network-Based hand posture control of a humanoid Robot Hand. IEEE Int. Conf. on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), 124-128, 2014.
28. Soderstrom, N.C. and Bjork, R.A. Learning Versus Performance: An Integrative Review. Perspectives on Psychological Science, vol. 10(2) 176–199, 2015.
29. Santos, O.C., Boticario, J.G. Practical guidelines for designing and evaluating educationally oriented recommendations. Computers & Education, 81: 354-374, 2015.
30. De Roeck, D., Slegers, K., Criel, J., Godon, M., Claeys, L., Kilpi, K., and Jacobs, A. I would DiYSE for it!: a manifesto for do-it-yourself internet-of-things creation. In Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design (NordiCHI '12). ACM, New York, NY, USA, 170-179, 2012.
31. Banzi, M. Getting started with Arduino. O'Reilly Media, 2009.
32. Santos, O.C., Boticario, J.G. Exploring Arduino for Building Educational Context-Aware Recommender Systems that Deliver Affective Recommendations in Social Ubiquitous Networking Environments. Web-Age Information Management. LNCS, vol. 8597, 272-286, 2014.
33. Santos, O.C., Saneiro, M., Boticario, J.G., Rodriguez-Sanchez, C. Toward Interactive Context-Aware Affective Educational Recommendations in Computer Assisted Language Learning. In New Review of Hypermedia and Multimedia, 2015 (in press).

Realizing the Potential of AIED

Lewis Johnson

Alelo Inc.

Abstract. This is a time of opportunity and promise for AIED as a field. The field has had some major successes, and is having an impact with significant numbers of learners. Now that Big Data has arrived in education, opportunities are opening up to generate analytics from that data and use it to personalize learning. There is however potential to have an even greater impact on education, and make greater use of AI technologies. The field should focus on realizing this potential, and not divorce itself from either AI or Ed. Achieving impact will require more effective dialog and collaboration with educators, learners, and people in industry.

Keywords: Educational impact, partnering with education, partnering with industry, participatory design, technology transfer

1 The Time for AIED has arrived

These are exciting times for learning technologies. Technology is becoming integrated into education at all levels, as online learning, blended learning, and smart classrooms are becoming the norm. The global market in technology-enabled learning is projected to grow at an annualized rate of 20.3% to \$220bn in 2017 (MarketsandMarkets, 2014). Large as this is it is still a small fraction of the \$5.89tr that is projected to be spent on education in 2015 (Next Up Research, 2010); this suggests there will be even greater opportunities in the future. Technology-enabled education is enabling and fuelling demand for personalized and adaptive learning and assessment (Borden, 2011; Getting Smart, 2012), capabilities which AIED systems are well positioned to provide.

AIED-based systems are contributing to this innovation in learning. Alelo's language and culture training systems (Johnson, 2010; Camacho et al., 2009; Johnson et al., 2012) are in widespread use throughout the world, with well over 100,000 learners to date. They have had a significant effect on the cultural and linguistic competence of the learners who use them. For example the 3rd Battalion, 7th Marines, the first American Marine unit in the Iraq war to complete their tour of duty without any combat fatalities, learned Iraqi Arabic language and culture using Alelo's Tactical Iraqi learning environment (Marine Corps Center for Lessons Learned, 2008). Another AIED success story is the ASSISTments system, which is being used throughout the United States by nearly 20,000 or more students per year (Gelfand, 2011). And perhaps the biggest success so far has been the Carnegie Learning curriculum and software, which as of 2010 had been used by over 500,000 students (Institute of Education Sciences, 2010).

The workshop call for papers questions whether the ideas of AIEd are influencing AI or Education in any major way. The above examples illustrate that it is AIED is in fact having an impact. One could perhaps argue as to whether they are having a major impact, but they certainly intend to do so.

Yet these examples are just the beginning, and AIED has the potential to have an even greater impact on education in the future. The challenge for the AIED community is to realize that potential. It needs more success stories – examples of AIED research that is having an impact. The more instances there are of research that is having an impact, the more impact the field as a whole will have.

I regret that other obligations do not permit me to participate in person in the workshop in Madrid. However remote participation is becoming commonplace in technology-enabled learning, so I hope it is also possible for a major international conference on technology-enabled learning such as AIED. In any case I feel compelled to contribute this position paper and hopefully offer some constructive suggestions.

2 Connect AIED to Educational Problems

I have a number comments on the questions posed in the call for papers, but I will focus here on just one: the extent to which the results of AIED research are meaningful to real educational practices. Or to put it another way: What steps can people in the AIED community take to ensure that their research has meaningful educational impact? Here are some recommendations.

Talk with educational leaders. More than individual teachers, educational leaders and managers have a broad view of how where the unmet educational needs are, and may be open to innovative approaches that can meet those needs. Many of these are needs that AIED technologies can address. If you have a promising AIED technology, show it educational leaders and listen to what they have to say. They might help you make the connection to education needs, or if not you will come away with a better understanding of what the critical educational needs really are. They may be able to put you in touch with schools and teachers that are receptive to innovative solutions.

Talk with people in the edtech industry. There is not enough dialogue between AIED researchers and people in the edtech industry, which leads me to suspect that that there may be an insufficient appreciation of what researchers can learn from such dialogue. People in edtech have an understanding of what it takes to make a real impact on real educational problems with technology. They may be aware of educational problems that they themselves are not in a position to address, but they wish someone else would.

Engage in effective iterative, participatory design. The workshop call for papers suggests that participatory research is often a matter of rhetoric rather than practice. The question as I see it is how to make such participatory research achieve more effective results. Dialogue with educational leaders prior to the start of the design process can help, to make sure that the design is focusing on the right problems. So can iterative participatory design, in which researchers show teachers and learners partial prototypes and ask for input on how to improve it. Participatory design can be very effective when people have something concrete to respond to.

Learn from research programs that value educational contributions. The US National Science Foundation's Cyberlearning program is an example of research program whose projects address learning research questions as well as learning technology questions. The program requires research teams to carefully evaluate the educational impact of the designs that they develop, instead of simply focusing on technology development. Other AIED researchers can draw useful lessons from this and similar programs.

The RALL-E project (Alelo, 2015) is an example of an exploratory AIED research project that has undertaken each of these steps. With funding from the National Science Foundation's Cyberlearning program, we have developed a lifelike robot that can converse in Chinese, using the Robokind's Zeno-R25 robot as a platform. We developed the concept with advice from the Virginia Department of Education, which made us aware of critical needs in their state such as the lack of availability of qualified language teachers in many schools and the lack of access to high-quality interactive learning materials in many of those schools. We designed RALL-E as an interactive language-learning tool that students can use to develop their conversational skills, with or without the presence of a teacher. The Virginia Department of Education introduced us to the principal of a receptive test site, the Thomas Jefferson High School for Science and Technology (TJ) in Alexandria, Virginia. We have developed the robot iteratively, and have conducted a series of focus group tests with students and teachers at TJ. This has helped us refine the technical concept, as well as develop a better understanding of how it might be used in an educational context. This gives us confidence that students and teachers will respond positively to the completed solution. And finally, we talk with other people in the edtech industry, to determine how this technology might be relevant to educational needs that they see.

As more AIED projects draw lessons from projects that have had good impact, it will help the field overall to realize its potential of improve education. The rapid increase in availability of computing resources is multiplying the opportunities for the field to make a difference. If we seize these opportunities the prospects for the future of AIED are bright indeed.

Acknowledgments

This research was supported in part by the National Science Foundation under Grant IIS-1321056. Any opinions, findings, and conclusions expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

References

1. Alelo (2015). Alelo Develops an Interactive Robot for Learning Chinese. Retrieved March 16, 2015 from http://www.prweb.com/releases/chinese/robot/prweb12538156.htm.

2. Borden, J. (2011). The Future of Online Learning. *eLearn Magazine*, August 2011. Retrieved March 16, 2015 from http://elearnmag.acm.org/featured.cfm?aid=2024704.
3. Camacho, J., Johnson, W.L., Valente, A., & Bushika, M. (2009). Cultural training: The Web's newest gaming frontier. In Proceedings of I/ITSEC 2009.
4. Gelfand, A. (2011). Unleashing the Potential. *Worcester Polytechnic University Annual Research Magazine*. Retrieved March 16, 2015 from http://www.wpi.edu/research/magazine/unleashing.html.
5. Getting Smart (2012). New Survey Shows Demand for Personalized & Adaptive Learning Assessments. Retrieved March 16, 2015 from http://gettingsmart.com/2012/02/new-survey-shows-demand-for-personalized-adaptive-learning-assessments/.
6. Institute of Education Sciences (2010). Intervention: Carnegie Learning Curricula and Cognitive Tutor® Software. Retrieved March 16, 2015 from http://ies.ed.gov/ncee/wwc/reports/hs_math/cog_tutor/info.asphttp://ies.ed.gov/ncee/wwc/reports/hs_math/cog_tutor/info.asp.
7. Johnson, W.L. (2010). Serious use of a serious game for language learning. Int. J. of Artificial Intelligence in Ed. 20(2), 175-195.
8. Johnson, W.L., Friedland, L., Watson, A.M., & Surface, E.A. (2012). The art and science of developing intercultural competence. In P.J. Durlach & A.M. Lesgold (Eds.), Adaptive Technologies for Training and Education, 261-285. New York: Cambridge University Press.
9. Marine Corps Center for Lessons Learned (MCCLL) (2008). “Tactical Iraqi Language and Culture Training System“ *Marine Corps Center for Lessons Learned Newsletter* 4 (8), 4.
10. MarketsandMarkets.com (2014). Smart Education & Learning Market by Hardware (IWB & SBL), Software (LMS/LCMS, Open Source & Mobile Education Apps), Educational Content (Digital Content, Test And Assessment & Digital Text Book) - Global Advancements, Market Forecast and Analysis (2014 - 2019). Retrieved March 16, 2015 from http://www.marketsandmarkets.com/Market-Reports/smart-digital-education-market-571.html.
11. Next Up Research (2010). NeXt Knowledge Factbook 2010. Retrieved March 16, 2015 from www.kellogg.northwestern.edu/faculty/jones-ben/htm/NextKnowledgeFactbook2010.pdf.