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)