27. The great challenges and opportunities of the next 20 years

Each section of this chapter is written by a different author focusing on a particular challenge or opportunity for the future of their own choosing. There are various ways in which the sections could have been organised but in the end, we left them in the order that they were originally submitted. Note however that Sections 1, 3, 7 and 9 all deal with issues around *access and equity*, Sections 4, 5, 8, 10 and 11 cover different kinds of *learning and enabling technology* and Sections 2 and 6 offer two aspects of the same issue, namely *ubiquity and information overload.*

1. AIED AND EQUITY

Maria Mercedes T. Rodrigo

The United Nations Sustainable Development Goal #4 aims to "ensure inclusive and equitable quality education and promote lifelong opportunities for all" (UNESCO, 2021). The problem this goal seeks to address, though, is massive and dire: 262 million children are out of school, two-thirds of children do not acquire basic literacy and numeracy skills despite several years in school, and 750 million adults are illiterate.

AI-based learning technologies offer a partial solution. Intelligent tutoring systems (ITSs), adaptive instructional systems and other similar technologies have been shown to improve learning outcomes, outperforming other instructional methods such as traditional classroom instruction, reading printed or computer-based texts and laboratory and homework assignments (Streenbergen-Hu & Cooper, 2014). Indeed, the effect sizes of computer-based intelligent tutors approach those of human one-on-one tutoring (VanLehn, 2011).

AI-based technologies are particularly promising for low-prior-knowledge (LPK) learners. A meta-analysis from Ma and colleagues (Ma et al., 2014) showed that, regardless of prior knowledge, learners achieve comparable learning gains from ITS use. Indeed, ITSs offer LPK learners the opportunity to catch up with their high-prior-knowledge (HPK) peers (Jackson et al., 2010). LPK learners benefit more than HPK learners from immediate feedback (Razzaq et al., 2020) and from polite rather than direct intelligent tutors (McLaren et al., 2011). As the technology matures even further, some researchers speculate that we could design AI-based educational interventions to make use of eye trackers, brain signals, cameras and microphones to monitor student attention and dynamically redirect it in order to make better use of sparse attention resources (D'Mello, 2016). Some researchers are already integrating intelligent learning technologies with Massive Open Online Courses (MOOCs) in order to make adaptive learning systems available at scale (Aleven et al., 2018).

Under-resourced classrooms, whether in the Global North or South, seem like the environment in which AI-based educational systems can achieve the most impact. These technologies offer learners who are lagging behind a chance to raise their levels of achievement and potentially their life outcomes. However, many factors prevent this from taking place. Of these, we discuss three: accessibility, cultural bias and teacher preparedness.

Lack of access to hardware and the internet is the most obvious impediment to use of ITSs and other AI-based learning systems. Five percent of US families with children in the United States do not have computers and 15% have no internet subscriptions (Holstein & Doroudi, 2021). Many low- to middle-income families are under-connected, that is, they have slow or mobile-only internet access, or family members have to share devices. Some researchers have suggested that educational interventions be designed specifically for use on mobile phones, the primary computing platform of the Global South (Nye, 2015). However, people from lower-income groups tend to use low-cost internet access, buying internet time and bandwidth as needed and using it conservatively (Uy-Tioco, 2019). Telecommunication companies provide zero-rated internet that allows users access to certain news and social media sites. These plans, though, limit the types of content and media that subscribers can access. AI-based interventions, even when deployed on mobile technologies, are unlikely to be part of these zero-rated plans and will tend to be beyond reach as these technologies are often built to run on 5G networks (Gallagher, 2019).

A second impediment is cultural bias. Most ITS and AI in education research takes place in Western, educated, industrialized, rich and democratic (WEIRD) countries (Blanchard, 2012). In terms of language alone, many AI-based materials are made for English speakers, hence adoption in non-English-speaking countries requires extensive translation (Holstein & Doroudi, 2021). Beneath these surface features, the models themselves on which these systems base their intelligence are created using data from WEIRD populations and do not reliably transfer to other cultures. Ignoring cultural and contextual factors can decrease a technology intervention's effectiveness on a target group (Baker et al., 2019). The study by Ogan et al. (2015), for example, illustrates how most ITSs are designed for one-on-one instruction and do not work as intended when used in Chilean classrooms where students often work collaboratively.

A third impediment is teacher preparedness and support. Many teachers, especially those in the Global South, lack the training and exposure that they need to effectively integrate AI-based interventions in their classes (Rodrigo, 2021). Many have experience with productivity tools such as word processors and spreadsheets, but they do not have the skills to make use of these and other resources in innovative ways. Teachers lack in-service training opportunities to improve on their skills. They also lack curriculum support: often, there is no pre-made curriculum that integrates an ITS's use. Teachers do not have a lesson plan that guides their use of an innovation in a class or even an articulation of the alignment between the ITS's learning objectives and curriculum goals (Nye, 2014).

The result of these and other impediments is a "Matthew Effect", also known as an accumulated advantage phenomenon: those who already have access to sophisticated educational systems will continue to improve their knowledge, skills and life opportunities (Gallagher, 2019). This pattern is already evident in studies regarding the use of MOOCs. MOOCs were seen as democratizing technologies because they could provide quality education at little to no cost, therefore increasing the number of people with post-secondary education (Ebben & Murphy, 2014; Veletsianos & Shepherdson, 2016). Summaries of MOOC learner demographics, however, show the opposite: those who benefit are those who are already privileged, that is, people aged between 20 and 40 who already have a post-secondary degree (Ebben & Murphy, 2014; Gynther, 2016). Those who are less privileged will lag behind even further, unless course-corrected by other social, political or economic forces. Good education and good educational technologies are among these forces.

How, then, can ITSs and other AI-based interventions contribute to greater equity? We ask a few questions that might point to ways forward, in relation to the impediments cited earlier. In terms of access: How can researchers and developers design AI-based interventions for lower technology platforms? Instead of assuming high-end computers or mobile devices and high-speed internet access, can we capitalize on more broadly available but less sophisticated devices and platforms? In terms of cultural bias: How can we work with stakeholders from the Global South to create systems that address their cultures and contexts? How do we adopt pedagogies, content and teaching/learning strategies so that they embrace a greater diversity of learners and their circumstances? Finally, in terms of teacher preparedness: How can we create mappings between our systems and curricula? Can we build in teacher training, monitoring and support as part and parcel of our research undertakings? If we have constructive answers to these questions, we stand to increase the adoption and impact of the technologies we create and move all our learners towards the realization of universal, inclusive, equitable and high-quality education.

2. ENGAGING LEARNERS IN THE AGE OF INFORMATION OVERLOAD

Julita Vassileva

AI-powered tools free us from unpleasant, annoying, or boring tasks and make our lives more comfortable and easier. Yet intelligent technologies also make us lazier, both physically and cognitively. The abundance of information, including online learning sites and applications, that are available at the tip of our fingers is both a blessing and a curse. We register the new information, but do not have time to make sense of it, to see how it can be integrated with what we already know; we live in a constant "cognitive dissonance". Thinking, understanding and aligning knowledge structures requires time and focused attention – a resource we lack, due to our fear of missing out on something new in the meanwhile. We keep scrolling down to find information that we already know and that confirms our current beliefs and prejudices; this makes us feel comfortable and safe, more in control of the information barrage. How to know which information or authority (source) to trust in an age when old authorities are questioned, old monuments are falling to the ground, and political corruption is exposed on a massive scale around the world? We are paralyzed by hesitation – what is worth remembering or learning?

Nevertheless, humanity does adapt to the technologies it creates. An increasing proportion of human skills are automated and we lose them. Do we even need to be literate? Do we need typing or handwriting skills? We now dictate text to the computer, smartphone, smartwatch or laptop, and they produce text that can be easily processed. Who needs spelling or grammar skills? There are excellent spellcheckers, and decent grammar- and style-proofing tools. Duolingo has revolutionized learning languages and made it available to everyone everywhere, in an engaging, playful way. But why even bother to learn a new language? Computer translations are already quite good, at least between the most common languages. Google translates for free even specialized legal documents, which normally require an experienced professional translator in the area.

Professions and entire categories of work are transforming or even disappearing at an unprecedented rate; the profession "translator" (at least from languages such as English, French, Spanish and German) will almost certainly employ very few people in the future. Similarly, many professions requiring high-level training, e.g., family doctors, lawyers, financial advisors, possibly civil engineers and architects will likely be significantly reduced in numbers. It could happen to teachers as well. We have already experienced telemedicine and telelearning on a massive scale during the pandemic and it seems to work relatively well, while yielding massive efficiencies. All these jobs require skills and knowledge developed after years of systematic training, and they are likely to disappear or become extremely competitive. For what jobs should we prepare our children? What skills should we learn to be able to requalify ourselves after losing our jobs, when big changes happen in such a short time? Is human learning becoming obsolete? Is there a place for education, training, and mentoring in the future?

I believe that we need to commit a lot more attention to human learning, just as we do now to climate change, and support it with technology to meet the new challenges. It seems that basic literacy, numeracy, and language skills are important "Lego blocks" needed to build higher-order (meta-) skills for abstract thinking, categorization, and organization, as well as discipline, grit, and ability to focus attention on a problem and not give up. It is hard to learn high-level skills directly. Skills for problem solving, discovery, creativity, planning, and decision making are best learned in the context of specific domains, and learners generalize their experiences through a lot of practice and failures. New meta-skills will be needed by the lifelong learner in the age of information glut: skills for searching and evaluating the credibility of information sources, communication and collaboration skills, skills of meditation to focus the mind, fight procrastination, and the patience and determination required to study and practice challenging concepts and skills.

Young people now communicate less face-to-face and more online. This is not only due to the pandemic waves with the recurring school closures and masking regulations, but also because young people are growing up being baby-sat by smart phones and tablets, and social networks entice them early on into non-stop communication with their peers through screens. Yet it is not clear that this communication leads to learning social and collaboration skills, or how to empathize with others. On the contrary, we see an increase in mental health issues at increasingly younger ages. It becomes important also to learn "hygge" or how to "switch off" technology, to keep peace of mind, to appreciate poetry and deep friendships. With all of these categories of knowledge that seem necessary (basic skills, higher-level thinking skills, information-handling skills and mental health skills), what is the minimal basic common knowledge and skills that young people need to learn to be able to function in a society?

These are hard problems that cannot be solved by intelligent computers; they need to be solved by society, by humans. Yet, AI-powered technologies can provide educational tools that help to learn some of the skills listed above. Many AI-powered educational tools and apps are already freely available online and their number will grow constantly. They offer personalized learning paths to acquire specialized knowledge in many domains, for example, astronomy, biology, a language, or food chemistry. Learners will use these tools to develop domain-specific skills in areas of their choice, so that they can build upon them and practice their higher-order (meta-)skills: discovery, critical thinking, problem solving, and decision making. Can we use AI-powered technology to teach meta-skills? As the recent misinformation epidemic shows, we really need intelligent tutors to teach skills in critical and selective information consumption, problem solving with basic probability and statistics. We need tutors using Socratic dialogues challenging learners' existing beliefs and stereotypes, and helping them expand their knowledge, views, and online social circles. There are already persuasive technology apps that advise people on how to maintain their physical and mental health, and AI-based systems that mentor users on how to work effectively together with others, respecting others' autonomy, views, and cultural background (Dimitrova & Mitrovic, 2021). Further research should explore how to better personalize these systems using holistic learner models and to use data mining with different data modalities: not just text, but also voice, motion and gesture, facial expressions, and various context features to uncover the cognitive, affective, and motivational states, and the current goals of the learner.

While the big unsolved problems described above are related to the question "What to teach?", many new problems also arise around the question "How to teach?". The advances in learner modelling need to be matched by advances in the pedagogy for personalized instruction and guidance. For example, how to use the insights into the learner's mental state to select the best teaching or motivational action? Machine learning can help discover successful pedagogical interventions for identified learner states, if there is a lot of training data available. This leads to the need to share learner data widely and openly among developers of AI-based educational systems. Balancing the need for sharing data with privacy legislation and security constraints will be challenging. Using anonymized data-sets and synthetic data to pre-train models can alleviate privacy concerns, but may increase bias and inequality, if not done carefully.

One problem that will be gaining importance is how to maintain the learner in a favourable motivation state for learning. This is especially important when they are trying to acquire conceptually difficult knowledge or to practice skills that require a lot of mundane and repetitive training. While there has been significant progress in detecting learner attention and affect using eye-gaze and various sensors, discovering motivation states is not straightforward. How to influence a deteriorating level of motivation with appropriate actions? A human teacher may choose to address the student's decreasing motivation by many means, for example, by letting the learner try a simpler problem to boost their self-efficacy, by cracking a joke, by emphasizing the importance of a particular concept or skill relative to the overall learning, by reminding and reinforcing the learner's commitment to their learning goal, or by suggesting to take a break. But how to discover whether the decreasing motivation is the short-term result of negative affect, such as a temporary distraction, fatigue, or because of a long-term change in goals, interests, and priorities? How to proactively engage the student in a dialogue to disambiguate the diagnosis, without becoming obtrusive? A lot of relevant research on conversational agents has been done over the last 25 years in specific domains and learning contexts; as natural language generation technology and conversational agents are becoming ubiquitous (Alexa, Google Assistant, Siri), the education field is ripe for applying these technologies widely and developing new ones.

In addition to maintaining learner motivation during a particular learning experience, an important problem is getting learners motivated to embark on learning something. As technologies evolve and take over tasks that were done by humans, we all need to be life-long learners. How to keep motivated to invest the time and effort to learn a new domain of knowledge and skills, when it is not certain how long before computers take over and the knowledge/skill also becomes obsolete? The optimistic view is that people will always want to learn new things, because they are naturally curious. The pessimistic view is that people will "give up" and resign themselves to the state of passive consumers, sustained by universal income, seeking satisfaction in a meta-world, playing games or taking drugs that make them feel happy and fulfilled.

Educational games and gamification of learning systems are ways to make learning more engaging and attractive. There has been a lot of research in this area over the last ten years which has shown strong positive results, and yet still relatively few approaches use personalization and adapt the game mechanics to the learner's engagement, motivation and knowledge levels (Rodrigues et al., 2021). There is untapped potential in using machine learning and data mining in the context of educational games and gamification. Dynamic personalization of gamified educational systems and educational games is an open research area that is likely to attract a lot more attention in the future.

Persuasive technology is an area of research that can offer useful strategies and insights into increasing user motivation and engagement. There is a large set of persuasive strategies that can be used to influence user behaviour without coercion, for goals that benefit the user and society. Until recently this area has developed applications mostly to stimulate behaviour change in e-commerce or in health-related apps, helping users to eat properly and exercise regularly. There is an active research stream in personalizing persuasive strategies considering psychological types, culture, and demographic factors (Oyibo et al., 2017), personality features, goals, and context. Attempts to incorporate dynamic personalization by learning from user response to persuasion are ongoing. Applications of personalized persuasive technologies are now also appearing in educational systems (Orji, 2021; Orji & Vassileva, 2021), and mental health (Adib et al., 2021; MacLeod et al., 2021) and this is a promising direction of research to address learner engagement, and, more generally, to promote behaviours that are conductive to acquiring higher-level personal and interpersonal skills.

To summarize, the ongoing information revolution raises many new problems for education. Some of them relate to the question "how to teach and engage" and they can be foreseeably addressed with solutions relying on rapidly developing AI technologies, such as data mining, advanced learner modelling of affect with personalized games and persuasive technologies. However, there are many important problems that society needs to tackle and define, before technological optimizations can be found. They are related to the question "what to teach" and generally, how to make human learning socially rewarding again.

3. PEDAGOGICAL AGENTS FOR ALL: DESIGNING VIRTUAL CHARACTERS FOR INCLUSION AND DIVERSITY IN STEM

H. Chad Lane

Introduction

A pedagogical agent (PA) can be defined as a virtual character or physical robot that seeks to promote learning, enhance motivation, and provide support to engage in an educational activity. PAs can be non-interactive or interactive, and typically seek to emulate naturalistic communication with learners through speech, gesture, emotions, and action. In AIED systems, PAs are typically considered as part of the user interface and designed in the context of system features intended to support learning, such as for problem solving and with natural interaction (see Chapter 10 by Lajoie & Li, Chapter 11 by Rus et al.).

There are many reasons to think that PAs can be beneficial. Given that much learning is fundamentally social and occurs with other people, such as in classrooms (Kumpulainen &

Wray, 2003) and museums (Leinhardt et al., 2003), it is intuitive to extend AIED systems to include naturalistic interactions including questions, dialogue, smiling, pointing, and more. It is established that people are prone to adopt social frames when interacting with computers (Reeves & Nass, 1996) and so including an embodied PA may activate an even deeper social frame for learners as they gain new skills or knowledge.

Over the last three decades, substantial effort has gone into investigating the designs, roles and impacts of pedagogical agents (Heidig & Clarebout, 2011; Johnson & Lester, 2018; Lane & Schroeder, 2022; Moreno et al., 2000). The field has worked diligently to identify evidence-based design principles (Baylor, 2011; Craig & Schroeder, 2018) and to conduct wide-ranging efficacy studies (Schroeder et al., 2013). So the challenge now lies in how best to leverage this important technology as it inches closer to becoming more mainstream and having the potential for greater societal impact. In this contribution, I argue that the next phase of research should focus on the power of PAs to promote more inclusive and welcoming learning environments. Specifically, I describe a roadmap that positions PAs as *ambassadors of STEM education* and that focus on improving learner feelings about domain content. We should view PAs as tools in an ongoing battle to bring equity and diversity to STEM (Ferrini-Mundy, 2013). Operationally, I advocate for investigation of identifying strategies that seek to invite learners to participate, engage, and persist in STEM such that their learning experiences have longer-lasting effects that sustain the goal of the acquisition of knowledge as essential, but also produce meaningful impacts on noncognitive outcomes, such as interest, persistence, academic resilience, and motivation.

The Critical Influences of Appearance and Social Interaction

A long list of factors influence the relationship children have with STEM. First, children are bombarded with depictions of STEM professionals on television, in advertisements, and even in the video games they play. Decades of research has demonstrated that these (often subtle) messages shape attitudes and beliefs about their own self-concept and identity (Dill-Shackleford et al., 2017). Second, a child's relationship with STEM is highly influenced by the social STEM experiences in their lives. For example, the presence of "everyday science talk" around friends and family (Cian et al., 2021), access to early, informal STEM learning experiences (Dou et al., 2019), and classroom activities with teachers and peers (Hachey, 2020) are all known to play important roles.

PAs can be considered to address both of these aspects directly. For example, the decision to use a PA is an implicit agreement to also design a representative of the topic being taught.[1] Early research indeed confirmed that PA designs could invoke stereotypes by the learner (Moreno et al., 2002). Baylor and Kim (2004) explored some of the design space of agents, such as ethnicity and gender, concluding (among other findings) that agents that challenged stereotypes lead to greater transfer of learning. However, this finding has proven difficult to replicate. For example, Schroeder and Adesope (2015) found that gender had no impact on cognitive or affective outcomes in their studies. Continued investigation into the subtle consequences of PA design is much needed.

Perhaps more importantly, however, PAs that engage in natural interaction with the learner (e.g, spoken or typed input, natural language output, or nonverbal signals) are aligned with the second key influencer of STEM identity: social experiences. Some PA research is suggestive that fine-tuning social behaviors can fundamentally change a learning experience. For example, Finkelstein et al. (2013) found that a peer agent that used a *culturally congruent* dialect (i.e., the same dialect as the learner) produced improved science communication and science reasoning by learners. This study suggests that, for younger learners, PAs should seek to communicate in familiar ways that enable learners to focus on expressing STEM understanding and less on whether or not they are speaking "correctly". A different study on agent behaviors showed that a PA in an informal learning environment who conveyed excitement and high levels of interest in the topic produced fewer non-productive learning behaviors and greater feelings of self-efficacy in learners (Lane et al., 2013). Many more studies support the idea that agent behaviors can influence both cognitive and affective outcomes (Schroeder & Adesope, 2014), an essential point if we adopt the goal to position PAs in more prominent roles in the future.

Roadmap for Leveraging Pedagogical Agents for STEM Inclusion

Although the design of PAs for more "traditional" learning environments is far from solved, there are reasons to assume PAs can achieve much more and have more lasting impacts. In this final section, I outline three goals for future research that emphasize their potential roles in addressing challenges associated with broadening participation in STEM (NSF INCLUDES Coordination Hub, 2020).

Support for social and emotional learning (SEL)

A key barrier for participation in STEM is that learners from historically underrepresented and underserved groups often face stress and anxiety from numerous sources. These can be related to legacy inequalities, such as lack of access to needed resources (e.g., internet), or to factors related to poverty. While a PA is not likely to solve such systemic problems, awareness of such factors and the ability to convey empathy could be part of a coordinated effort (say with educators or counselors) to promote the emotional self-regulatory skills, like academic resilience and persistence, that would enable a young disadvantaged learner to become more engaged with STEM. Arming PAs with knowledge about how to bolster SEL skills holds incredible promise to lead to downstream success in STEM (Garner et al., 2018).

Entry points to STEM and support for interest development

Early learning experiences are critical in the development of STEM learning (McClure et al., 2017) and for long-term interest development in STEM (Renninger & Hidi, 2016). For example, young learners in early stages of interest typically need to experience more positive emotions and require high levels of support to receive positive feedback and encouragement (Renninger et al., 2019). While ideally such support would come naturally through peers and adults, it doesn't always happen. A PA that is specifically designed to gently introduce a topic to a learner, use personalized motivational techniques (e.g., that reflect personal relevance), and offer ongoing encouragement could have a positive effect on a child (just as their television and media consumption do).

Identify learner assets and broaden adaptivity of PAs

Finally, I propose that AIED researchers seek a new breed of PAs that are able to exercise a new level of adaptivity. While beyond the scope of current technologies, it is certainly not unreasonable to imagine a PA being able to infer and elicit more about the learners they interact with. Specifically, *asset-based* approaches to STEM education suggest that deficit-based approaches (i.e., correcting what is "wrong" about learners who are not in STEM) can be equally as damaging as systematic societal biases. Such a PA would be more driven to learn about the learner, their interests, their strengths, and their possible goals. This would then directly influence the pedagogical choices that are made, such as what content to present and feedback policies.

An example content area for such an approach can be found in the (now numerous) STEM programs in the US that build on Black youth interest in Hip-hop (Emdin et al., 2016). It is possible to imagine a PA adopting the goal of "bringing out the STEM" in such domains and building learner confidence to learn and achieve self-selected goals. Highlighting connections to STEM, such as between computational thinking and music, or digital skills for music production, would be essential. The design of a knowledge base for culturally aligned learning activities would be needed, but strengthened with the capability for dynamic retrieval of content from the web to improve relevance and timeliness (Zhai & Massung, 2016).

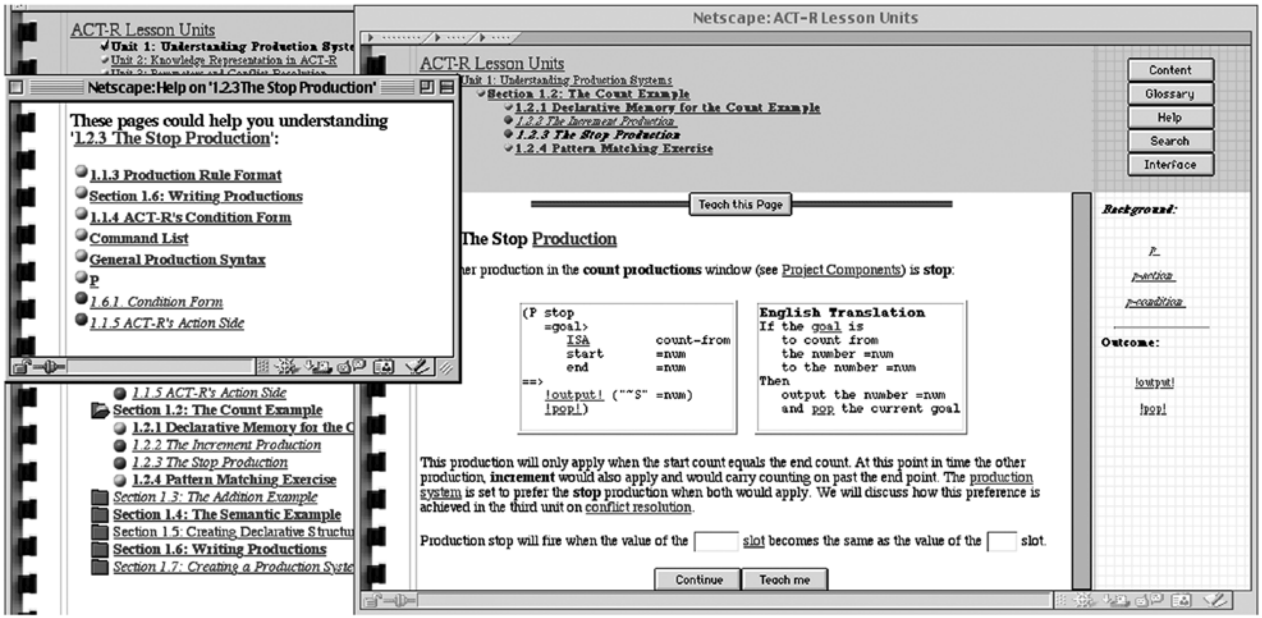
Conclusion

Substantial progress has been made in PA research over the last 30 years, demonstrating their efficacy to both enhance learning and define fundamental design principles. The trajectory of this broad body of work along with rapid advances in relevant areas of AI suggests that a new wave of PAs is coming. It is not clear whether current PAs tap fully into their full potential power to create social learning experiences, and so a new generation of PAs that specifically address goals related to inclusion and cultural adaptivity could have profound impacts on how learners view their own learning experiences. I have proposed that this new generation of PAs should include techniques to support the social and emotional learning needs of children, focus on "ambassador"-like skills such as introducing STEM domains and promoting interest, and gaining a new level of adaptivity to support learning in ways that adapt to emergent learner goals. Together, these goals could position PAs as an important tool in creating AIED systems that are more inclusive and which catalyze other related efforts to broaden participation in STEM.

4. INTELLIGENT TEXTBOOKS

Peter Brusilovsky and Sergey Sosnovsky

Textbooks are one of the most popular tools for learning. Traditional textbooks are a product of many decades of refinement and improvement. Page numbers, table of contents, index, and glossary are examples of important empowering functionalities that were added over the years and are now expected by every reader. However, a quiet revolution in the area of textbooks that we are witnessing now is going far beyond past attempts on improving textbooks. Starting in the early days of multimedia with the first interactive textbooks on CD-ROM, every new technology wave encouraged publishers and authors to explore new opportunities for textbook production and distribution. With their popularity and support growing year by year, digital textbooks are gradually replacing their printed counterparts. In turn, it encourages researchers from several fields to explore digital textbooks as an application area for advanced technologies. Over the last 25 years, the research on *intelligent textbooks*, that is, the use of AI to



Source: Brusilovsky, 1998

Figure 27.1 *Adaptive navigation support and content recommendation in an adaptive textbook developed with InterBook* enrich textbooks, has gradually become one of the main directions of textbook innovation. Below, we review the research on intelligent textbooks by examining opportunities provided by several popular AI technologies.

Crafting Textbooks Augmented with Intelligence

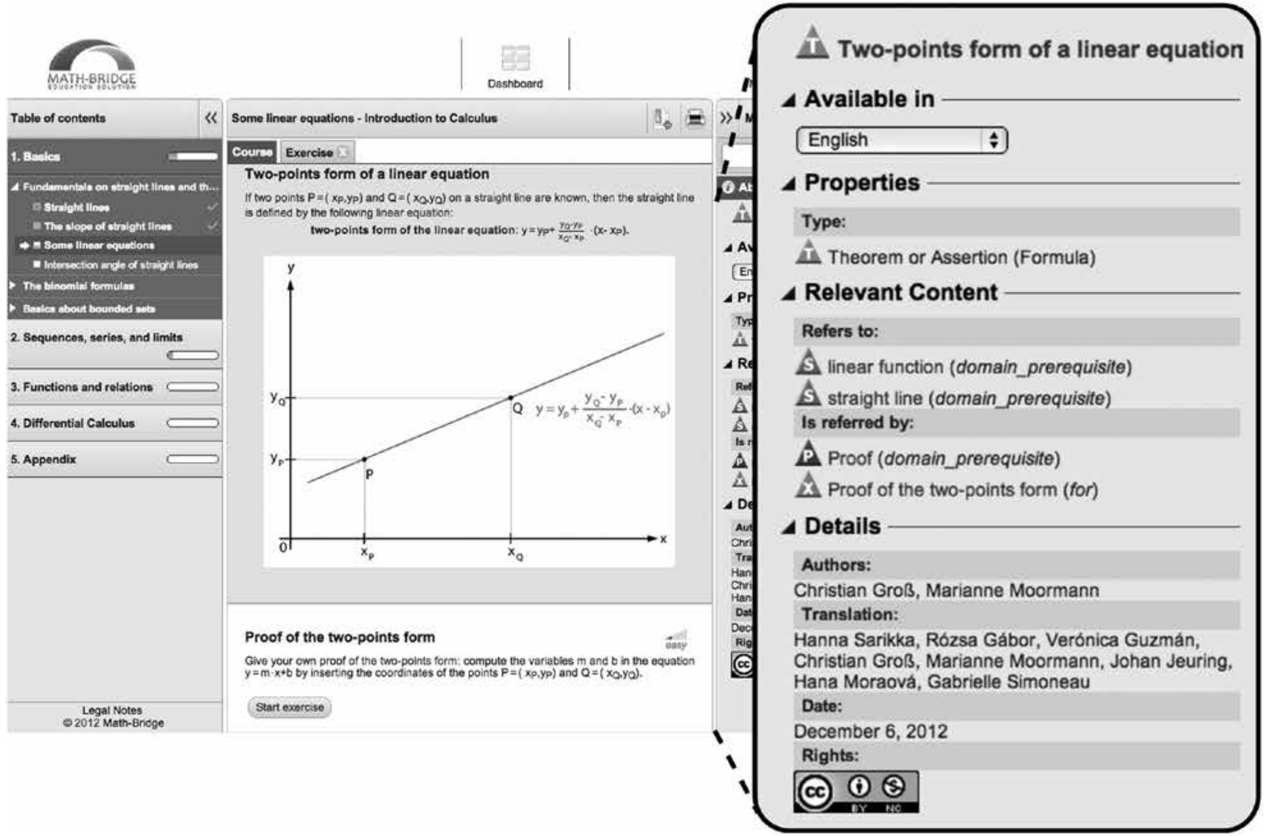
The first wave of intelligent textbooks was developed by researchers on *adaptive hypermedia* within the traditional stream of AI research focused on knowledge elicitation and representation (Hohl, 1996; Brusilovsky, 1998; Weber, 2001). The goal of adaptive hypermedia textbooks was to guide different readers to the most relevant sections of hypermedia-based textbooks using *adaptive navigation support* or to choose the most relevant content to present through *adaptive presentation*. The intelligence and personalization in these books were based on augmenting textbook sections with domain knowledge elicited from domain experts. For every textbook, domain experts were expected to create a *domain model* (a set of key concepts presented in the book) and use it to create a *content model* for every book section (i.e., a set of concepts covered by the section). Most advanced adaptive textbooks used domain models in a form of semantic network and distinguished prerequisite/outcome concepts for each section (Brusilovsky, 1998). The knowledge of individual learners was represented as a traditional *overlay student model* which maintained an estimation of student knowledge for each domain concept. Early adaptive textbooks demonstrated that these knowledge models could support a number of intelligent functionalities – such as various kinds of adaptive link annotation (Brusilovsky, 1998; Weber, 2001), adaptive presentation (Hohl, 1996), content recommendation (Kavcic, 2004), concept-based navigation (Brusilovsky, 1998) and textbook assembly (Melis, 2001).

Enriching Textbooks with Intelligence

One of the classic problems of early intelligent textbooks was that they followed a closed-box approach. Once implemented, it was impossible to modify them or enrich them with additional content or knowledge. The development of Semantic Web technologies offered new standardization formats and modeling languages for shareable and expressive knowledge representation along with new architectural solutions for intelligent software. This allowed intelligent textbook researchers to implement knowledge models as ontologies and educational material as interlinked and annotated learning objects, facilitating integration of external content. Initial projects developing these ideas focused on proposing new fully ontological architectures that allowed manual extension of adaptive textbooks with external learning objects (Dolog et al., 2004). Later, successful attempts were also made to automate semantic integration of external content into adaptive textbooks (Sosnovsky, 2012) and to adaptively assemble textbooks from large repositories of interlinked learning objects (see Figure 27.2). Semantic Web has also supplied intelligent textbook researchers with the expressive power of description logic. When combined with symbolic natural language processing (NLP), it supported development of semantically annotated textbook prototypes that can “understand” their content and engage in rich interactions with students, such as meaningful question answering and concept mapping exercises (Chadhuri, 2013).

Extracting Intelligence from Textbook Content

Over the last decade, a few projects started investigating textbooks from a very different perspective. They looked at the textbooks not as an object to add value to, but rather as a source



Source: Sosnovsky, 2014

Figure 27.2 *Interface of the Math-Bridge system: an adaptively assembled textbook with links between different learning objects displayed in the right-hand panel* of value. This new generation of research has been enabled by advances in the fields of NLP and machine learning (ML). The content, the structure, and the formatting elements of textbooks have been utilized to extract domain knowledge and different types of learning objects. A number of papers explored approaches for automatic extraction of topics and concepts from textbooks (Wang et al., 2016), as well as harvesting relationships between them (Adorni, 2019). Ultimately, this line of research leads to automatic construction of full concept-based domain models "behind pages" (Alpizar-Chacon, 2021). These models can be used to support intelligent and adaptive services pioneered by the early textbooks (such as concept-based navigation or integration of external content), yet without expensive manual knowledge engineering. Another application of NLP technologies that has gained prominence in the last few years is the use of textbooks to generate additional learning content, such as assessment questions (Kumar, 2015) or definitions (Yarbro, 2021).

Extracting Intelligence from Textbook Interaction Data

Several interesting directions of research on intelligent textbooks have been motivated by rapidly increased volume of data capturing user interactions with online textbooks. Existing research leverages these data in several ways. Most importantly, large volumes of learner data enabled researchers to better understand reading behavior and connect it to learning progress. On a coarse-grained level, understanding patterns in textbook navigation and annotation

Source: Barria-Pineda, 2019

Figure 27.3 Social comparison and external content recommendation in ReadingMirror behavior could be used to predict learner success or failure, which could be used for early intervention (Winchell, 2018; Yin, 2019). On a fine-grained level, student interaction with a textbook could be used to model student knowledge for different textbook concepts. These fine-grained models could be used to recommend most relevant internal (Thaker, 2020) or external (see Figure 27.3) content to read.

User interaction data could be also used to improve the textbook organization or representation. For example, in combination with content data, log data could be used for a more reliable elicitation of prerequisite relationships (Liu, 2016). Finally, in the spirit of transparency, interaction data could be released directly to the readers through the textbook interface. This approach could be used to support open learner modeling and social awareness. Figure 27.3 shows an example of a socially transparent textbook interface ReadingMirror (Barria-Pineda, 2019) where a combination of personal and social reading data offers social navigation and social comparison.

5. AI-EMPOWERED OPEN-ENDED LEARNING ENVIRONMENTS IN STEM DOMAINS

Gautam Biswas

Introduction

The last twenty years have produced significant developments in AI, machine learning and technologies that support data collection and analyses in computer-based learning environments. New directions have evolved in the field of AI in Education, one primarily being the development of Open-Ended Learning Environments (OELEs) that adopt a constructivist epistemology to support the acquisition of domain knowledge along with critical thinking and problem-solving skills (Biswas et al., 2016). Students working in these environments have a specified learning goal, such as constructing a model of a scientific process. To facilitate learning with understanding and to help students develop problem-solving skills (Bransford et al., 2000), we have adopted AI and machine-learning methods that provide students with tools and resources that support hypothesis generation, solution construction and hypothesis verification and refinement in different phases of learning. Betty's Brain (Biswas et al., 2016), C2STEM (Hutchins et al., 2020a) and SPICE (Zhang et al., 2019) are examples of OELEs developed by our group. In addition, we have developed machine-learning methods and learning analytics to analyze students' learning behaviors and performance from their activity logs in our OELEs.

AI Representations to Support STEM Learning

In all of our systems, we adopt AI-driven modeling representations (e.g., causal maps, block-structured programming languages) that are intuitive, visual, and executable. The ability to execute their evolving models helps students develop important reasoning, explanation, and debugging strategies (Biswas et al., 2016; Zhang et al., 2021). In recent work, we have exemplified the importance of AI representations in developing domain-specific modeling languages (DSMLs) to scaffold students' computational modeling of scientific phenomena (Hutchins et al., 2020b). We illustrate the use of AI representations in two example systems below.

The *Betty's Brain* learning environment (Biswas et al., 2016) presents students with the task of teaching a virtual agent generically named Betty, about a scientific process by constructing a visual causal map. The DSML representation of the causal map provides a set of concepts relevant to the science topic and directed links that students can use to model causal relations between concepts. Betty uses a simple qualitative reasoning mechanism with the map she has been taught (Leelawong & Biswas, 2008) to answer questions and explain her answers. The student's goal in this OELE is to teach Betty a causal map that matches a hidden, expert model of the domain. The students' learning and teaching tasks are organized around three primary activities: (1) reading hypertext resources, (2) building the map, and (3) assessing the correctness of the map. The hypertext resources describe the science topic under study. As students read the resources, they need to identify causal relations and then explicitly teach those relations to Betty in the form of the causal map. To check the correctness of their current map, students can ask Betty to take a quiz. The quiz results help students determine the correctness and completeness of their current map. Figure 27.4 illustrates the Betty's Brain system quiz interface.

SPICE (Science Projects Integrating Computing and Engineering) implements a novel middle- school three-week NGSS-aligned water runoff curriculum (WRC) (McElhaney et al., 2020; Zhang et al., 2021). Students take on an engineering design challenge to redesign their schoolyard and overcome flooding problems during periods of heavy rainfall. The challenge is to minimize surface water runoff during a storm, while adhering to constraints on the total cost, accessibility of the schoolyard, and the need to accommodate different play activities (e.g., soccer field and basketball court). The science investigation involves hands-on activities and developing a conceptual model that focuses on the *conservation of matter*, that is, the amount of water runoff is the difference between the total rainfall and water absorbed by surface materials. Students are then introduced to

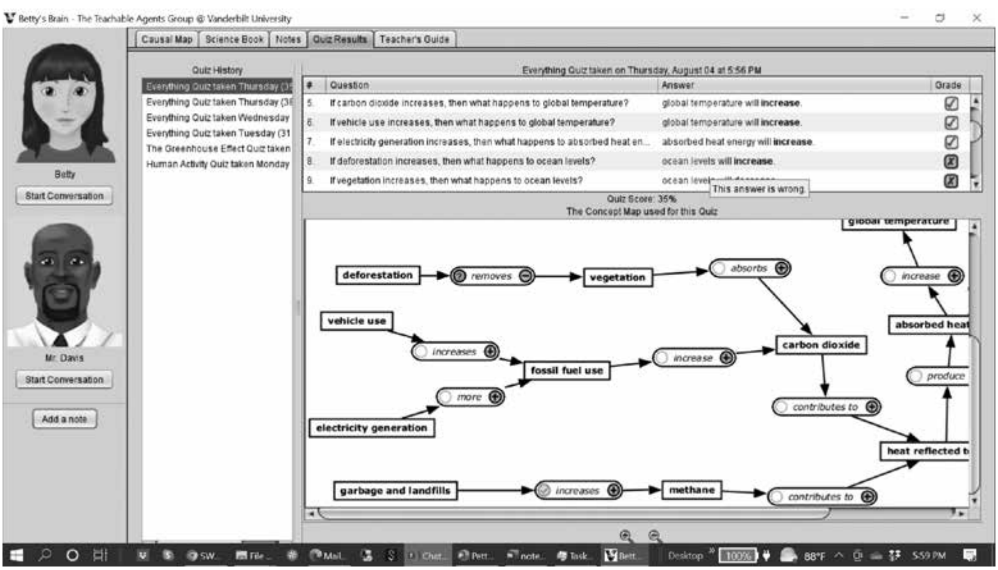


Figure 27.4 *Quiz interface in the Betty's Brain system*

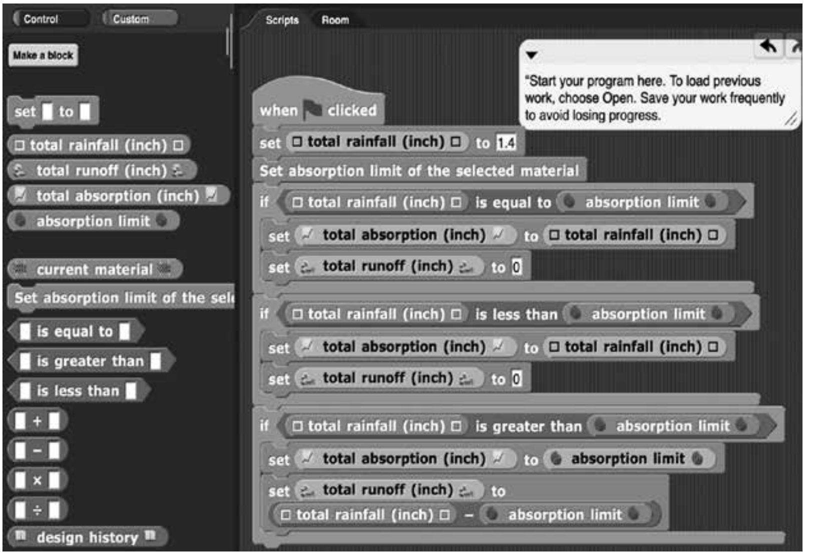


Figure 27.5 Computational Modeling Interface in the SPICE environment

the computational concepts of variables, expressions and conditionals, which they apply to develop a computational water runoff model. Like before, this model is executable and supports testing and debugging activities to help students refine their models. Figure 27.5 shows the DSML created for the computational modeling activity (left) and a correct implementation of the runoff model (right). Students then use their runoff model code in an environment comprising 4×4 squares representing the playground layout for the engineering design task. Students select appropriate materials for the different parts of the schoolyard. The model, when executed, computes the total run off and cost for the current design solution.

These examples illustrate the use of AI representational and reasoning mechanisms in our OELEs:

- The use of domain-specific, intuitive, visual representations mechanisms that make explicit the concepts and reasoning mechanisms of the science domain (Hutchins et al., 2020b);
- Providing students with a sequence of linked domain representations (e.g., conceptual, computational, and engineering design models in SPICE (Hutchins et al., 2021) that facilitate learning with multiple representations;
- An exploratory, inquiry-based approach to learning of scientific processes using executable models, so students can formulate and test hypotheses, and learn to verify (debug) their models, thus developing critical thinking skills.

Analyzing Students' Learning Behaviors

When working in OELEs, novice learners often have difficulties in keeping track of the tools provided in the learning environment and combining the use of these tools in an effective manner to accomplish their goals (Basu et al., 2016). These problems are exacerbated by the students' lack of self-judgment and strategic thinking abilities. To understand students' challenges and help them make progress in their learning tasks, we have developed machine-learning methods to track students' performance and learning behaviors as they progress through a sequence of curricular units (Kinnebrew et al., 2017; Hutchins et al., 2021). Our approaches combine model-driven learning analytics and data-driven sequence mining methods in machine learning to understand students' difficulties (Kinnebrew et al., 2017).

For example, we have developed sequence mining methods (Kinnebrew et al., 2013) to analyze students' activity data capture in log files to extract their frequent activity patterns and associate them with learning behaviors and strategies (Kinnebrew et al., 2017). In addition, we have developed analytic measures based on the concept of coherence analysis (Segedy et al., 2015) to understand the relations between their consecutive actions in the context of their associated tasks. In studies with Betty's Brain conducted in middle-school classrooms, a frequent pattern we have observed students performing is: *Add Incorrect Link* (AIL) → *Take Quiz* (Quiz) → *Remove Incorrect Link* (RIL). Students add an incorrect link to their map, follow that by taking a quiz, and then make the deduction that the link added was incorrect and remove it from their map. A first reaction is that this pattern represents a *Guess and check strategy* used by the students. However, a deeper analysis reveals that *high-performing students*, that is, those who succeed in building a correct causal model, tend to use this activity pattern mainly toward the end of their map building tasks. On the other hand, students who had difficulty in building correct models use this pattern frequently when they commence their map-building tasks.

This indicates differences in the use of the pattern by the two groups of students. To investigate further, we computed two analytic measures: (1) *pattern coherence*, implying the link added was the link removed; and (2) *pattern support*, implying the quiz results provided evidence that the link removed was incorrect. These results show that the high performers were more likely (3:1) to use this pattern for systematically checking the link they had just added (pattern was coherent and supported), whereas the low performers were much more likely (10:1) to use this pattern as an uninformed guess and check strategy (their use of the pattern was not coherent and not supported).

As a second example, we investigate how students' performance in each one of the curricular activities in SPICE influenced performance in subsequent activities. We developed a number of measures that included pre- and post-tests that covered science, engineering, and CT to quantify students' prior knowledge and knowledge gained through the intervention, respectively; formative assessments to scaffold student learning, students' performance in building their computational models, and the quality of their engineering design solutions generated.

We used a machine-learning approach, *Path Analysis* (Ahn, 2002), to study the relative importance of effects among the measured performance values for the sequence from pre-test to formative assessment to computational modeling to engineering design to post-test scores in the WRC curriculum. Path analysis represents a form of structured equation modeling (Kline, 2015), and may be considered as a simplified form of learning a Bayesian network (Pearl & Mackenzie, 2018). Figure 27.6 shows the derived path model.

The figure points to a number of informative pathways that form the basis of understanding students' progress through the curriculum. For example, higher prior knowledge in computational thinking led to better performance in the formative assessments, then better

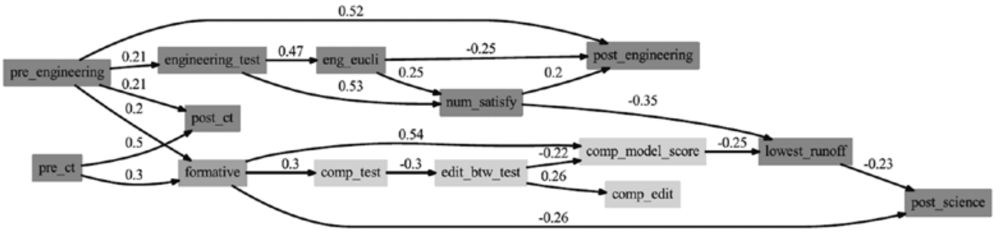


Figure 27.6 *Causal paths discovered with statistically significant direct effects*

computational model-building strategies (e.g., building the model in parts, testing the model for correctness), which in turn led to the building of more correct computational models of the science phenomenon. Further, this path showed that these students also generated the best design solutions, that is, solutions that met the cost and accessibility constraints and minimized runoff. Last, they had higher scores in their science and CT post-tests. Similarly, higher prior knowledge in engineering led to using more effective search methods and testing to generate design solutions, resulting in more optimal design solutions as well as higher science and engineering post-test scores.

In summary, the two examples above, demonstrate the use of machine-learning techniques that we have employed to analyze logs of students' activities in our systems to gain a deeper understanding of their learning behaviors and their performance in learning STEM concepts and practices.

The Future: From Log File Analysis to Multimodal Analytics to Understand Students' SRL Behaviors

Researchers have established comprehensive frameworks for studying self-regulated learning (SRL) as a dynamic, interacting collection of students' cognitive, affective, metacognitive, and motivational (CAMM) processes (Azevedo et al., 2017; Panadero, 2017). Our prior work primarily focused on understanding students' cognitive and metacognitive processes and strategies. However, recent work has used additional sensors to better understand students' learning behaviors. For example, we have used:

- *Microphones* and *natural language processing* techniques to capture and interpret students' conversations and understand the difficulties they face in their computational modeling tasks when they work in pairs in the C2STEM environment (e.g., Snyder et al., 2019);
- *Eye-tracking devices* and machine-learning *classifiers* to predict how students' gaze patterns during reading resources are predictive of their abilities to construct correct causal maps in Betty's Brain (Rajendran et al., 2018); and
- *Face-tracking devices* and *deep-learning* algorithms to understand the relations between students' affective states and their cognitive and metacognitive behaviors (Munshi et al., 2018; Taub et al., 2021).

Currently, we are developing *multimodal analysis frameworks* that combine data from multiple sensors, such as students' activity logs, microphones, videos, eye-tracking, and face-tracking devices, along with advanced machine-learning algorithms to build comprehensive models of learner behaviors and performance as they work in OELEs. Our long-term research goals are to design adaptive scaffolding in OELEs to help students overcome their learning difficulties and become self-regulated and independent learners.

Acknowledgments

This research is partially supported by NSF (IIS) Cyberlearning award # 2017000 and a previous NSF STEM+C award DRL-1742195. The author would also like to support a long list of graduate students and collaborators who have worked with him on a variety of research projects in the area of AI in Education.

6. UBIQUITOUS-AIED: PERVASIVE AI LEARNING TECHNOLOGIES

James C. Lester

AI has emerged as a powerful family of technologies that promise to reshape every facet of society. AI systems are rapidly becoming more accurate, training times for machine learning are falling dramatically, and the range of AI-driven applications is expanding at an unprecedented rate. Through "edge AI," natural language processing, computer vision, and machine learning are being deployed on a broad array of platforms that operate in the field, not just in the cloud. Perhaps the most striking development in the emergence of AI is its ubiquity. With the proliferation of ever more powerful computing devices, sensors, and networks that support extraordinary volumes of data, AI is quickly becoming deeply enmeshed in our everyday lives.

Just as AI is rapidly becoming ubiquitous, so will AI learning technologies. This development introduces both the opportunity and the need to re-envision how learners will interact with AI-empowered learning environments. A significant departure from how the pioneers of AIED imagined the systems they were designing would be used, *ubiquitous-AIED* calls for a research program that fundamentally changes how we envision the field. It pushes to the forefront issues of the learner operating in a wide range of naturalistic learning contexts.

Assumptions of Ubiquitous AIED

AI capabilities are becoming increasingly commoditized with core AI technologies quickly becoming exceptionally accurate, fast, and inexpensive. Because AI-driven learning environments directly leverage AI technologies, hardware and software barriers to widescale usage of AIED are rapidly falling. In a very few years, it seems all but certain that AI-driven learning environments could be physically distributed throughout schools, informal learning contexts (e.g., museums, after-school programs), homes, and workplaces, and running on devices with a range of form factors from the large (e.g., wall-sized displays) to the mid-sized (e.g., laptops) to the small (e.g., phones, watches). They could also be delivered through increasingly powerful media that include virtual reality, augmented reality, and mixed-reality media. With the proliferation of increasingly powerful and inexpensive sensors, they could have access to many modalities of learner data spanning auditory (speech) and visual (learner facial expression, gaze, gesture, posture) data streams, as well as locational and physiological data streams.

In addition, they could readily have access to temporal data about learners spanning windows of milliseconds, minutes, hours, weeks, months, and even years. Whether they *should* have access to these data is a deeply important question.

Ubiquitous-AIED Opportunities

When AI becomes a utility, AIED can become pervasive, and the prospect of ubiquitous AIED opens up a wide vista of AI-empowered learning. Learner models will become unimaginably powerful. In addition to the high-volume streams of multimodal data for an individual learner, learner models will be trained on data from a multitude of learner populations in myriad learning contexts for countless subject matters and skills. They will achieve levels of accuracy never previously imagined. Their capabilities for recognizing learners' goals and plans will grow rapidly, and the range of the competencies they model will dramatically expand. In addition to cognitive and disciplinary competencies typically associated with traditional curricula, they will be able to accurately model metacognitive and social skills, complemented by highly accurate affective models. Significant increases in diagnostic power will shape individualization of learning experiences in ways that we cannot currently envision. They will be able to generate problem-solving scenarios, scaffold learning, and create extraordinarily motivating experiences that are deftly crafted for individual learners in very specific learning contexts. Furthermore, because the contexts of AIED deployments will be represented at high levels of granularity, they will enable group learner modeling that supports rich collaborative learning interactions.

Ubiquitous-AIED will profoundly reshape AI-empowered learning experiences. Consider the case of pedagogical agents. Learners will engage with ubiquitous-AIED pedagogical agents that are unfettered by current-day computational constraints. Learners will converse with their agents in rich spoken dialogue, regardless of whether the agent is on their laptop, phone, or watch. Agents will be able to reason about learners' non-verbal signals as learners express delight, furrow their brow, use a gesture to convey nuance, or walk with a quickened gait.

Agents will tailor problem-solving advice not only to the learner herself but also to the physical context in which she is situated. And because agents will be able to draw on potentially years of experience interacting with the learner in many contexts, they will be able to create motivating learning experiences and provide guidance that are extraordinarily attuned to the needs and interests of that particular learner.

Ubiquitous-AIED will manifest in many ways. In addition to personified pedagogical agents with visual personae, it will drive speech-only pedagogical agents. It will unobtrusively shape learning experiences in text (e.g., learner-unique textbooks), audio (e.g., learner-unique podcasts) and video (learner-unique animations and documentaries). It will create educational gameplay experiences in which the levels and non-player character interactions are specifically tailored to that learner and learning context (including the platform). It will support multi-context learning in which learners move from one context, such as their elementary school classroom, to another, such as a museum or their home. It will seamlessly shift from coaching traditional subject matters to coaching sports to coaching health behaviors (e.g., nutrition) for wellbeing.

Ubiquitous-AIED will play a prominent role in the workplace. It will soon be hard to conceive of a worker performing a job in the absence of AIED technologies. Workers in sectors ranging from science and technology to finance, transportation, retail, media, agriculture, and healthcare will experience continuous learning with AIED technologies. The education workforce will experience a seismic shift: with the appearance of omnipresent AIED teaching assistants, K-12 teachers and university faculty will soon have at their disposal the means to create learning experiences that fundamentally change how instructors perform their jobs. And rather than experiencing artificially separated phases of working and training, workers will access their AIED technologies to acquire new competencies on demand. Furthermore, as AI–human "teaming," in which job responsibilities are shared between workers and their AI tools, emerges as the dominant paradigm, AIED technologies will take on a new role of scaffolding "teaming" competencies.

The arrival of ubiquitous-AIED will spawn a new era of learning analytics and assessment. The enormous volumes of learner interaction data noted above will yield learning analytics that provide deep insight into learning for researchers and educators. These in turn will give rise to a new generation of powerful assessment technologies for both formative and summative (including high-stakes) assessment. Ubiquitous-AIED-based assessment will be psychometrically grounded and achieve levels of reliability and validity never seen before. Notably, it will support measurement for a broad range of competencies and will be conducted in naturalistic contexts.

Design and Policy Implications for Ubiquitous-AIED

Emerging AI capabilities point toward design and policy implications for ubiquitous-AIED learning technologies. First, considerations of context become increasingly important. Because learners will interact with ubiquitous-AIED learning technologies in specific contexts, these technologies should be designed to reason about the cultural, social, and physical contexts in which they will be deployed. Thus, rather than the comparatively impoverished sense of context that AIED as a field has typically explored, it will be important to give primacy to considerations of the cultural, social, and physical contexts in which learning will occur, as well as their affordances for different aspects of learning. For example, free-choice learning environments such as museums suggest that AI learning environments should be designed to support short dwell times with large groups of learners interacting collaboratively, while home-based learning environments might be designed for much longer learning scenarios.

Second, it is imperative that we consider issues of AI ethics in the design of ubiquitous-AIED learning environments. Fairness, accountability, and transparency all emerge as matters of great importance because of the dramatically increased ability to affect learners' lives. Privacy concerns loom large. With rapidly rising capabilities for reasoning about individual learners, it is paramount that national and international policy be developed around considerations for creating ubiquitous-AIED environments that foreground diversity, equity and inclusion and have a firm ethical foundation that supports and protects all learners.

Third, AIED designers will need to imaginatively lift their sights to most effectively leverage powerful new AI capabilities. For example, with the new reality of natural language processing (NLP), computer vision (CV), and machine learning (ML) models growing nonlinearly in speed and accuracy, it will be important to investigate ubiquitous-AIED frameworks for pedagogical planners, learner modeling, and natural language tutorial dialogue that draw on underlying NLP, CV, and ML methods that support much more robust learning interactions than we have previously considered. And given the rate at which AI capabilities are increasing, we will need to design for multiple generations of AI beyond the current one, which introduces the opportunity to explore new learning technology paradigms supporting qualitatively new forms of learning.

Acknowledgments

This research is supported by the National Science Foundation under Cooperative Agreement DRL-2112635. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

7. CULTURE, ONTOLOGY AND LEARNER MODELING

Riichiro Mizoguchi

In this essay, I briefly discuss cultural, ontological, and learner modeling issues which I have been investigating recently, and then their implications for future research are discussed.

Cultural Aspects

Millions of different species exist depending on each other to enable sustainable development. Even at the individual level, people are so different in many respects that the differences sometimes cause misunderstandings or struggles. Nevertheless, we should be in a harmony with people. It is the reason why diversity has gained considerable attraction these days. The difference is caused not only by personal factors but by cultural factors. In the Asian culture, in the context of education, teachers are highly respected and have a rather strong initiative in the classroom, which contrasts with Western culture. The belief that what a teacher says is right so that students should accept it is not an extreme case in Asian culture. It is along the line that seniority is respected in Asian culture. At a concrete level, the meaning or connotation of color is different from culture to culture. Although red means danger in many cultures, it means happiness in China, and vermilion is a symbol of long life or a talisman in China and Japan, so shrines in China and Japan are painted in vermilion. This suggests that color coordination at the interface of learning systems needs careful consideration of such cultural differences.

Frankly speaking, these observations are not new but have been discussed in the papers presented at CATS (Culturally Aware Tutoring System) (CATS, 2015) workshops. Although culture has been investigated in sociology, anthropology and related fields extensively to date (Hofstede, 2022; Sociology Group, 2022), there exists a gap between the achievements and educational systems. In order to contribute to bridging the gap, I have built two ontologies of culture with my collaborators, Blanchard (2014) and Savard (2019). The former is designed based on the so-called declarative perspective and the latter is on the procedural perspective, the two being mutually complementary. These ontologies would be useful for the future design of culturally aware tutoring systems that have been discussed in CATS.

Ontological Aspects

The utility of ontology includes making the fundamental conceptual structure of complex entities explicit so that application systems can select necessary concepts from them and consistently exploit them. The above ontologies should evolve through intensive use and feedback.

It would be beneficial to mention another kind of utility of ontology application. What we have seen to date includes an ontology of some entities followed by its use in applications. I would like to introduce a completely different way of the use of ontology in which no ontology is built. It utilizes ontological thinking itself and is applied to English teaching (Alard & Mizoguchi, 2021).

Most of the conventional technology-supported learning systems employ technology for facilitating the learning/instructional process. On the other hand, the system we are talking about uses ontological theories in reorganizing the teaching material, in particular, that of the English grammar of tense and aspects that are notorious for the *ad-hoc* rules. Learners taught using conventional textbooks are often forced to memorize those rules without convincing reasons or justifications. To avoid such a memory-oriented instruction, Danielle Alard and I have tried to reorganize those rules by employing ontological thinking/theories about events. We eventually came up with the following ontological findings:

(1) An event is necessarily a past or future happening, and it cannot be a present happening because the present is an instant and an event needs a non-zero interval,
(2) A process is ongoing so that it is expressed according to the present progressive aspect, and
(3) There exists no event expressed by the simple present tense in the real world.

The third fact derives a very powerful rule that the simple present tense should be used to express what we call *floating* happenings by which we mean events that are not anchored at any definite time in the real world. Typical examples include "I go to school", "The sun rises from the East", and "Go down this street about 100 m, then turn to the right…" which are organized in rules for the use of the simple present tense such as the expression about *custom*, *general truth*, and *instruction* in the conventional textbooks. About ten rules are successfully reduced to just one rule that "Use the simple present tense to express floating happenings".

We went through all the rules about the tense and aspects extracted from several popular textbooks and investigated them in the way described above to come up with new rules whose number is about one-third of the current ones. What is good is not only the drastic reduction in the number of rules but also that most of them are comprehensible to students, which frees them from memorization of *ad-hoc* rules. Although some exceptional rules remain, the majority of the new rules are derived from the ontological nature of events and time so that students can comprehend them independently of their cultural background. Danielle Alard, who is also a teacher of English, has deployed our new rules in her class for more than three years and confirmed a remarkable improvement with surprisingly high satisfaction from the students. I can say this is a great success both for the learning of the English tense and aspects and the innovative use of ontology.

Learner Modeling Aspects

Learner modeling has been a dominant research topic in the AIED community for years because it is key to the adaptive behavior of the learning support systems (see Chapter 7 by Aleven et al. on domain and student modelling) (Nkambou, et al., 2010). Theoretically put, model-building of the learner's understanding of the learning material has been the main focus of the research of learner modeling, which has led to the research on how to represent the state of understanding of a learner supported by an academic interest in what understanding is. This is the reason why all the existing learner modeling methods are about representing how students *understand* the subject. People might have an objection that the bug model does not fit this observation. Although such an objection seems to make sense at first glance, it is not the case. The bug model takes care of misunderstanding which is still in the same realm as the modeling of understanding because it is another "understanding", just one which is incorrect. What we want to explore here is the modeling of "*not understanding*". Imagine a situation where a student is completely at a loss. She has no idea about what she should do next.

Note here that we do not intend to tackle an intractable task such as modeling how learners misunderstand or uncovering what kind of incorrect knowledge learners would have. The idea of modeling "not understanding" does not deal with learners' knowledge nor its use but deals with the search space in which knowledge is applied.

Let's employ the notion of the search space in GPS (General Problem Solver). Conventional learner modeling methods deal explicitly with knowledge (operators) and its applications to problem solving. They implicitly assume the existence of a search space within which learners perform problem-solving activities (searching for a solution) by applying the knowledge (operators). To represent the "not understanding" state of a learner, we explicitly deal with the search space itself rather than knowledge and its application. Such a student who is completely at a loss is represented as having no search space and a student who fails to find a solution despite her continuous effort might be because of her incomplete search space caused by a failure of establishing a correct search space rather than the lack of knowledge to apply. In particular, for helping low-performance learners who are often at a loss and need effective help to get themselves out of such a state, suggestions to establish an appropriate search space for the problem solving at hand might be more effective than the explanation of necessary knowledge to apply (Mizoguchi, 2020).

Implications for Future Research

Here, I summarize some implications derived from what has been discussed above.

Culturally aware systems

The reformulation of the achievements about culture and cultural differences in sociology and anthropology for its applications to education/learning would be beneficial for building culturally aware tutoring systems. Ontologies of culture or cultural differences would contribute to this enterprise in many ways because they provide the underlying conceptual structure of culture in a computer-understandable manner. Teaching cultural differences would be also a promising topic in which ontologies are useful for producing convincing teaching materials supported by an in-depth understanding of culture as well as cultural differences.

New way of the use of ontology engineering

The success of the new way of ontology use in the reformulation of the teaching materials themselves would open up an innovative way of applying ontology engineering to building learning support systems. The research of OMNIBUS/SMARTIES is viewed as cognate (Mizoguchi & Bourdeau, 2016). The OMNIBUS ontology is built to operationalize learning/instructional theories to build a theory-aware authoring system named SMARTIES. Those theories are descriptive and have to be interpreted by researchers to use them for building theory-justified learning support systems. OMNIBUS enables such manual interpretation to be skipped by providing operationalized theories that can be interpreted by the SMARTIES authoring system. The reorganization of the rules for the use of English language tense and aspects together with such operationalization of theories by the use of ontology should develop further in the future.

LM of "not understanding"

Although there remains much room for refinement of the method, it is a first step toward establishing a new paradigm of modeling the search space that has been ignored to date. In an extreme case, the notion of the search space itself can be a subject to learn. Not only students but ordinary people rarely have a clear notion of the search space. They tend to think or solve problems simply by applying the knowledge at hand without noticing that they cannot explore outside of the search space implicitly specified (limited) by the knowledge at hand, and hence they are sometimes at a loss when they fail to find a solution in a given time. In such a case, they should pay attention to the search space that has been implicitly assumed/specified and examine its possible limitations to extend or improve it by assessing the potential capacity of the search space. Note here that it is not the problem of whether or not learners possess relevant operators (knowledge), but rather the problem is of noticing relevant kinds of knowledge out of many they already possess to successfully extend the current search space.

8. CROWDSOURCING PAVES THE WAY FOR PERSONALIZED LEARNING

Ethan Prihar and Neil Heffernan

The desire for personalized learning comes from the idea that giving students instruction and support tailored to their specific needs or traits can lead to higher learning gains (Bulger, 2016). In the past, there have been statistically significant positive attempts to personalize learning, such as providing students with different instructional material based on their competency (Kulik & Kulik, 1982; Razzaq & Heffernan, 2009), and statistically insignificant attempts, such as giving students different material based on their assumed learning style (Pashler et al., 2008). Recently, the increased adoption of online learning platforms has enabled methods for personalizing students' learning to be incorporated into students' curricula at a much larger scale than what was previously possible in classroom settings, but while many online learning platforms purport to implement personalized learning, very few have evidence to support the efficacy of their methods (Bulger, 2016). The growing integration of online learning into students' curricula has provided online learning platforms with an unprecedented opportunity to improve their design by engaging with users to create and validate their content, as well as explore for further opportunities for personalization. Educators and researchers using these platforms can be solicited to create content for students, and the impact of this new content on students' learning can be evaluated empirically with data collected from students using the platforms.

ASSISTments, an online learning platform that has been running since 2006, not only has evidence that the platform is effective at increasing students' learning gains (Roschelle et al., 2016), but has also been engaging with educators and researchers to create tutoring for struggling students (Ostrow et al., 2017; Patikorn & Heffernan, 2020). Within *ASSISTments*, students are assigned mathematics problem sets from one of multiple open access curricula, which they complete online through the *ASSISTments Tutor*. The *ASSISTments Tutor* displays each mathematics problem for the students and offers on-demand tutoring if the student is struggling (Heffernan & Heffernan, 2014). Tutoring can come in the form of a complete explanation of how to solve the problem, individual steps that explain how to solve one part of the problem at a time, a series of simpler problems that walk the student through how to solve the more complicated problem, an explanation of how to solve a similar problem or feedback on the specific error a student made. Using the same terminology found in Chapter 9 by Aleven et al. on instructional approaches, *ASSISTments* implements the tutored problem-solving approach through *ASSISTments Tutor*, in which students are provided with step-level formative feedback. An example of the *ASSISTments Tutor* displaying a problem along with a text-based explanation is shown in Figure 27.7.

One way in which *ASSISTments* has leveraged its users to discover opportunities for personalization is by allowing external researchers to create and evaluate different tutoring strategies. Since 2015, more than 100 teams of researchers have been able to run randomized controlled experiments in *ASSISTments*. These experiments have compared video-based tutoring to text-based tutoring, examples of how to solve similar problems to examples of how to solve the problem the student is struggling with, language with positive sentiment to neutral sentiment, and more. By allowing these researchers to create experiments and evaluate different hypotheses, *ASSISTments* has collected over 50 000 instances of a student participating in an experiment. These data can be used to understand the relationships between traits of the students and the effectiveness of different tutoring. For example, one study in *ASSISTments* found that students with below-average prior knowledge benefited more from tutoring that broke the problem down into individual steps while students with

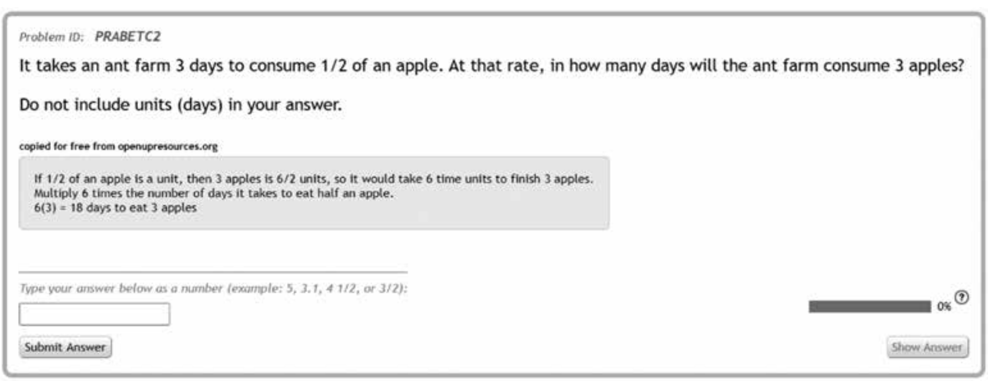


Figure 27.7 A problem in the ASSISTments Tutor with a text-based explanation

above-average prior knowledge benefited more from tutoring that gave them a full explanation of the solution all at once (Razzaq & Heffernan, 2009). By allowing external researchers to use *ASSISTments* as a tool for performing randomized controlled experiments, the insight gained from these experiments can be used to personalize content for students using the platform.

ASSISTments has also crowdsourced tutoring directly from teachers using the platform. Through *TeacherASSIST*, a program started in 2018, teachers in good standing that had created tutoring for their own students could agree to allow their content to be given to struggling students outside their class. A randomized controlled trial was conducted within the platform in which students that were struggling were given either just the answer or crowdsourced tutoring relevant to the problem they were struggling with. In 2020, the results of this study showed that the crowdsourced tutoring collected from teachers had a statistically significant benefit on students' learning gains (Patikorn & Heffernan, 2020). In 2021 this study was repeated and again the crowdsourced content was shown to have a statistically significant benefit on students' learning gains (Prihar et al., 2021). In addition to being able to determine the overall effectiveness of crowdsourced tutoring, the dataset collected on the effectiveness of *TeacherASSIST* was used to evaluate the individual effects of specific content creators' tutoring. It was discovered that while most content creators' tutoring was better than receiving just the answer, many of the content creators' tutoring was not statistically significantly better, and some content creators had content which was statistically significantly worse than just receiving the answer (Prihar et al., 2021).

Using crowdsourcing to collect tutoring not only helps struggling students but can also enable the exploration for opportunities to personalize students' learning. The *TeacherASSIST* project created an average of about three tutoring messages each for 6 044 problems in *ASSISTments*. For these problems where multiple tutoring messages are available, it is possible that one type of tutoring could be more beneficial to one group of students, and another type of tutoring could be better for a different group of students. This type of interaction, referred to as a qualitative interaction, is the ideal case for personalized learning because different content is more effective for different students. Using the crowdsourced tutoring and statistics collected on students and problems from prior usage data, it is possible to investigate qualitative interactions within the data collected through *TeacherASSIST*. However, with so much data available and so many potential interactions, using a statistical approach to test for qualitative interactions between the different tutoring messages available for each problem for every piece of information on students and problems would result in so many hypotheses that the correction for multiple hypothesis testing would take away any significance in the results. Therefore, to investigate qualitative interactions between students and crowdsourced content, *ASSISTments* has begun to use reinforcement learning to both provide the best content available to students and explore for opportunities to personalize students' tutoring. A reinforcement learning approach to search for opportunities to personalize students' learning would be able to learn over time which tutoring strategies are most effective for different groups of students and recommend the most effective content to students when they request tutoring. In simulations conducted on the *TeacherASSIST* data, even a simple Bernoulli Thompson sampling multi-armed bandit (Agrawal & Goyal, 2012) can increase students' next problem correctness by about 2.5% when selecting the tutoring the student will receive, and that doesn't even take into account specific traits of the student, problem, or tutoring. To explore the potential of reinforcement learning for personalization within ASSISTments, an algorithm, Decision Tree

Based Thompson Sampling, was designed to both personalize students' tutoring and identify qualitative interactions. This algorithm, by taking into account qualities of the students, problems, and tutoring strategies available, was able to increase students' next problem correctness by about 8.4%. In addition to increasing students' next problem correctness, Decision Tree Based Thompson Sampling was able to identify two qualitative interactions within the *TeacherASSIST* data. The first was that students benefited more from tutoring with images when the problem was more difficult than average, and from tutoring without images when the problem was easier than average. The second was that students benefited more from tutoring from a particular content creator when the problem was easier than average, and from tutoring from other content creators when the problem was more difficult than average. By crowdsourcing content from *ASSISTments* users, *ASSISTments* has been able to identify qualitative interactions and design algorithms that are key in ensuring that each student using the platform is given the content most helpful to them.

Crowdsourcing has paved the way for personalized learning within *ASSISTments* through the aggregation of step-level formative feedback (see Chapter 9 by Aleven et al. on instructional approaches), and can do so in other platforms as well. By encouraging users to create a variety of content, new algorithms capable of identifying interpretable qualitative interactions can be rapidly developed, tested, and deployed to discover new opportunities for personalization. Moving forward, platforms can even integrate public content as well as user-created content into the platform. By crowdsourcing opinions from teachers on publicly available educational videos, these videos could be offered to students when they are struggling with particular mathematics concepts, in addition to the tutoring already provided to them on the specific problem they are struggling with. Reinforcement learning algorithms can recommend what publicly available content and tutoring strategies the students should receive based on features of the student, the problem they are solving and the educational material available to them. These algorithms will not only determine the most effective tutoring for each student, but will also identify qualitative interactions that will provide insight into how to effectively personalize content for students, which can improve educational pedagogue and the design of future curricula.

Acknowledgments

We would like to thank NSF (e.g., 2118725, 2118904, 1950683, 1917808, 1931523, 1940236, 1917713, 1903304, 1822830, 1759229, 1724889, 1636782, & 1535428), IES (e.g., R305N210049, R305D210031, R305A170137, R305A170243, R305A180401, & R305A120125), GAANN (e.g., P200A180088 & P200A150306), EIR (U411B190024), ONR (N00014-18-1-2768) and Schmidt Futures for grant funding.

9. AIED IN DEVELOPING COUNTRIES: BREAKING SEVEN WEIRD ASSUMPTIONS IN THE GLOBAL LEARNING XPRIZE FIELD STUDY

Jack Mostow

Advances in education technology are enabling tremendous advances in learning at scale. However, modern educational software is typically developed in, by, and for WEIRD (Western, Educated, Industrialized, Rich, Democratic) countries, and therefore assumes resources taken for granted in such countries, including:

1) Reliable electricity,
2) Powerful computers,
3) High-bandwidth internet access,
4) Fast WiFi,
5) Accurate data collection,
6) Sophisticated sensors, and
7) Expert technical and instructional support to keep it all working.

Research to evaluate such software relies even more heavily on these assumptions due to the additional instrumentation and data collection involved. Though common in AIED, these assumptions are often violated in developing countries. This section discusses them in the context of the Global Learning XPRIZE (XPRIZE, 2019), a US$15 M worldwide competition to develop an open-source Android tablet app to teach basic reading, writing, and numeracy to children with little or no access to schools. XPRIZE tested the five finalists' Swahili apps in a 15-month-long independent controlled field study with over 2,000 children in 168 villages in Tanzania. An analysis of their learning gains appears elsewhere (McReynolds et al., 2020). This section examines how XPRIZE and some of the apps addressed each of the seven assumptions, whether by enabling it or by working around it.

Electrical Power

Educational apps typically assume reliable electricity. In developing countries, the reality tends to be quite different, with limited availability and fluctuating voltage.

The obvious and widespread workaround is to use batteries, but batteries must be replaced or recharged. XPRIZE worked around this limitation by providing a solar-powered station for each village with recharger cables, WiFi router, and ftp server. Another solution is to include a small solar recharger in or with each tablet.

Computers

Educational software in WEIRD countries can assume that computers are plentiful, powerful, and personal. XPRIZE enabled this assumption by providing a high-end tablet for every child at its study sites and replacing it if lost, stolen, or broken. Google donated 8,000 Pixel C tablets for the study. However, the reality is that computers in developing countries are typically scarce, low-end, and shared. Tablet sharing is not an issue for non-adaptive apps, but it is a problem for apps that depend on tracking individual progress to adapt to the student.

The *RoboTutor* team (Mostow, 2019; RoboTutorLLC, 2019) user-tested successive versions at several beta sites in Tanzania but could not afford to provide a separate tablet for each child. As a workaround, we supported tablet sharing by developing a self-enrollment and login mechanism that children could operate without being able to read, write, or require help from anyone who could. An alternative workaround developed by another team used continual assessment in its app to adapt quickly when the tablet changed hands. This approach eliminated enrollment but required adaptation based on scant data from the student.

Internet

Many educational apps assume reliable, wide-coverage, high-bandwidth internet access. The reality in developing countries is that internet access is limited, slow, or non-existent.

The obvious and necessary workaround was for apps to run without relying on an internet connection. Remote villages in the XPRIZE study had no internet connection. Cell phone service was sometimes available, but with bandwidth too low to transmit large quantities of data. XPRIZE's workaround was for its field staff to collect log data on portable devices and send it back via motorcycle to a location with a good internet connection. From there, XPRIZE staff uploaded each team's data to the Box cloud server for the team to download.

WiFi

Similarly, many apps assume reliable, wide-coverage, high-bandwidth WiFi. The reality in developing countries is that WiFi, when available at all, has frequent outages, limited coverage, and low bandwidth.

XPRIZE's workaround let tablets transfer files to and from the ftp server while within WiFi range, typically while recharging. The finalists' apps worked around limited WiFi access by running client-side without relying on the server, but took advantage of WiFi when available to transfer data, in particular log files.

Data Collection

AIED research in WEIRD countries can often rely on automated collection of comprehensive, accurate data. However, the reality in XPRIZE is that manual and automated data collection led to some incompleteness and incorrectness.

For example, whenever the Pixel C tablet battery ran out of power, recharging it reset the tablet clock to 12/31/1999, rendering logged timestamps invalid and wreaking havoc on analysis of log data. The *RoboTutor* team's workaround was to deduce lower and upper bounds on when they were actually created, based on the week that XPRIZE staff uploaded the log files. Another team's workaround to this problem was to advance the clock after each clock reset by a large, increasing time interval.

Fluctuating WiFi access apparently prevented transfer of an unknown number of files. A partial workaround for unknown data loss was to number log files sequentially and include their sequence numbers in their filenames. This remedy did not eliminate the data loss, but at least the sequence numbers made it possible to order log files chronologically and infer the number of missing logs from gaps in the log numbering sequence.

Field staff manually recorded a "tablet tracker table" listing the three-digit site ID assigned to each village, the six-character anonymized ID assigned to each child when pretested, the ten-character tablet IDs assigned to that child, and the dates assigned. The typos inevitable in manual data entry caused discrepancies in the form of "phantom tablet IDs" logged by the app but missing from the table. The *RoboTutor* team worked around such errors by finding whichever logged ID matched the table entry most closely.

Each week, field staff uploaded data to a folder named by the upload date, for example, 2018-12-22, with the subfolder for each site named by its site ID, such as 131. This process was manual and therefore vulnerable to "phantom village" errors. For example, we determined that the data in subfolder 131 of the 2018-12-22 upload folder actually came from site 139, based on the tablet IDs located there according to the tablet tracker table.

The lesson from these problems is that both manual errors (e.g. typos) and automated errors (e.g. clock resets) are inevitable, so redundancy in data collection is essential to work around them.

Sensors

Advanced educational software typically assumes that sensors are reliable, accurate, and many-modal, that is, exist for diverse forms of input (audio, vision, touch) and are used in a controlled environment. The reality in developing countries is that sensors are typically unreliable, inaccurate, few in number and used in noisy environments.

Various workarounds can help, for example:

1) Apply the sensor selectively, that is, only in situations where it can work. For instance, *RoboTutor* uses speech recognition only for highly predictable speech, such as listening to oral reading of a known text or listening for an expected answer to a problem.
2) Degrade gracefully. For instance, read a word aloud to the child after two attempts by the child to read it.
3) Provide a fallback. For instance, let the child tap to make *RoboTutor* read the next word.

Support

Educational software is typically used in environments where expert assistance and assessment – both technical and instructional – are available when needed. For example, a teacher or parent can help a student who has difficulty understanding the content or operating the software and assess whether it is at an appropriate level for the student. In developing countries, such expertise is often scarce or absent. Available adults are inexpert or even illiterate. Teachers are over-burdened, under-trained, absent, or non-existent.

Workarounds used by finalists' apps included verbally prompting the student what to do and pointing with an animated hand to indicate a target to look at or tap, because just highlighting it is not self-explanatory to users unfamiliar with touchscreens – the sort of cultural difference revealed by in-person study of the context of use. Such animations are useful within the context of an app, but video provides a useful workaround for unavailable human assistance by displaying external context. For instance, *RoboTutor* plays short video tutorials of humans demonstrating how to hold the tablet and operate various activities in order to work around the lack of someone to do it in person.

Finalists' apps used various workarounds to assess student knowledge in the absence of a human teacher. One simplistic workaround is to advance through the curriculum in a fixed order starting from the easiest lesson. This workaround relies on the unrealistic simplifying assumption that the resulting sequence will fit every student's rate of learning. Another workaround is to let students choose what to do. This "learner control" makes the unrealistic assumption that children's goal is to optimize their learning, and that they know which choice will do so. This assumption can be mitigated somewhat by restricting the choices to activities whose prerequisites the student has completed. A more sophisticated workaround is to automate assessment. Automated placement determines where in the curriculum the student should start. Automated promotion determines when the student has made sufficient progress to advance. Automated assessment uses and updates a student model.

Conclusion

The Global Learning XPRIZE was a landmark experiment in education technology. We hope that this discussion of AIED assumptions violated in developing countries may help AIED scale toward a world-class education for every child on the planet.

Acknowledgments

The Global Learning XPRIZE created the opportunity for this project. A large, dedicated, almost entirely volunteer team developed *RoboTutor*, with funding from generous donors and the $1M Finalist Award, and enthusiastic support from the Simon Initiative at Carnegie Mellon University.

10. THE FUTURE OF LEARNING ASSESSMENT

Claude Frasson

Traditional Assessment of Learning

Since Plato, Archimedes, and Pythagoras, assessment of learning has been the subject of many theories, methods, and practices. Thus, faced with behaviorism (Skinner), which considered learning as the association of a response to a stimulus, Piaget was able to develop constructivist approaches based on the idea that knowledge is constructed by the learner based on the mental activity. The learner is active and will seek to make sense of what he perceives from his experiences.

For Vygotsky, the learner builds his knowledge by socializing himself with fundamental tools such as language. Cognitivism aims to study observable behavior by ignoring mental representations that are not observable. This notion of the observable joins that known in quantum mechanics of an observable physical quantity (position, momentum, energy) capable of being broken down into quantifiable and deductible states. But how do you know if the learner has learned? Traditionally, pedagogical, psychological, or statistical theories are used to explain or deduce the knowledge that has been mastered and that which is not, from the performances obtained. But this constitutes an indirect assessment of the learner by his results and does not consider his mental state at the time. It is totally unknown whether the learner is tired, distracted, or not engaged in the learning process. Sometimes, the lack of interest in a teacher can cause difficulties in acquiring the subject. Educational approaches, involving the notion of learning style, have tried to provide solutions by taking more account of the individual. Thus, the "big five" theory (Golberg, 1981) made it possible to consider five types of learning relating to the personality traits of learners. There is therefore a progression of approaches tending towards a deeper knowledge of the individual, but this will remain insufficient until emotions are considered (Picard, 1997; Chapter 5 by Arroyo et al.) in learning.

Importance of Emotions in Learning

To come back to Archimedes, how did he discover the law of the principle of flotation? He was in his bath, in the ideal emotional conditions of relaxation and concentration to watch his leg float and suddenly he understood (Eureka!). The same is true of many fortuitous discoveries (Fleming, Pasteur) which have drawn conclusions from an observation, a high state of concentration and chance (phenomenon of serendipity). Indeed, emotions are narrowly linked to the cognitive process (Frasson & Heraz, 2011). Without emotions, we are unable to make even the simplest decision. Moreover, some strong emotions, like anger or anxiety, can prevent us from concentrating on our everyday tasks and particularly on learning (Goleman, 2006). People who have positive emotions will think in a more creative, expansive, and motivated way. On the other hand, negative emotions (stress, fear, disgust, boredom), can result in sleeping trouble, lack of concentration or hyperactivity and reduction of cognitive and memory functions.

Cognitive processes, such as problem solving and decision making, not only depend on the individual's emotional state, but are greatly intertwined with it (Chaouachi et al., 2010). Some emotions or combinations of emotions lead to a mental state which is in fact the condition of cognitive receptivity of the brain. Among these mental states, engagement and mental workload are fundamental. A series of failures can lead to a decrease in engagement but a series of successes will lower attention.

It has been found (Chaouachi et al., 2010) that for some learners, a consecutive series of wrong answers lowered their mental engagement. For others, the same effect was observed for a consecutive series of good answers. Indeed, learners tend to relax after getting a few good answers and consequently become less engaged. This is where the contribution of the emotional factor is crucial and must be assessed permanently to adopt adequate pedagogical strategies. Experiments have shown that emotions impact the engagement index, that the engagement index is more correlated with arousal than valence, that it represents a valid indicator of learners' performance, and construction of a personalized model to predict the variation of that engagement index is not only possible but highly recommended.

Mental workload is an indicator of acceptability, of the availability of the learner's mental resources (Berka, 2007). It is useless to explain anything to them if their mental workload is high (saturated); they will not be able to understand, hence the importance of being able to measure it. An experiment (Chaouachi et al., 2011) showed real-time monitoring of engagement and mental load. Faced with the impossibility of understanding a problem (too much workload), simple advice was enough to lower this workload and suddenly allow the problem to be understood.

Contribution of Neurosciences

Often researchers stick to their usual discipline to analyze or build experiments. However, the richness of different disciplines (and particularly the neurosciences) is that they constitute different points of view making it possible to find complementary solutions and original methods of investigation. Thus, the study of the components of the brain allows us to better understand the fundamental role of the cortex, the thalamus and especially the limbic system in the acquisition of knowledge. The latter regulates both the acquisition and the deployment of knowledge with the amygdala (center of emotions and implicit memory), the hippocampus (explicit contextual memory) and areas of short-term and long-term memory. Understanding this architecture of the brain is very important not only to understand the process of acquisition, memorization, but also to measure them. This time, we no longer consider indirect performance measures but take direct measures of the mental state of the learner by picking up brain signals. The most reliable and accurate physiological signal for monitoring cognitive state changes remains the electroencephalogram (EEG). The EEG data has a high level of time resolution and precision resulting from electrical neural activity in the brain. They measure four EEG bands: *Theta* (4–8 Hz), *Alpha* (8–13 Hz), *Beta* (13–32 Hz) and *Gamma* (32-40 Hz), which, by combination, allow assessment of various mental states.

In addition, neuroscience measures provide access to the distinction of brain areas involved in information processing. This is important for understanding how the reasoning is triggered. It is thus possible to distinguish processing by the conscious (left part, with the language area) and by the subconscious (right part, with the image processing area). We can follow the forms of reasoning which can switch from the deductive mode to the inductive mode which represents a great progress compared to the studies which sought to know by tests of approximations, the famous path of knowledge. A surprising experiment was made by Heraz et al. (2009), using the P-300 (positive) and N-300 (negative) brain waves: a learner had to choose a correct answer from three questions. When he chose the right answer, a P-300 brain wave was generated, but when he chose a false one it was an N-300 wave which manifested, which means that the subconscious part of the brain could guess that the answer was wrong.

Brain Assessment Parameters and Tools

As we have seen, the most important aspect of future learning assessment is to allow a direct assessment of the cognitive processing brain conditions of the learner and no longer a probable estimation of the reasons for their performance. EEG-based brain–computer interfaces use sensors placed on the head to detect brainwaves and allow mental states of fatigue, engagement, workload, stress, anxiety, concentration, or distraction to be distinguished. Great progress has been made in the precision and ease of use of these devices.

However, if you explain something to a learner while they are looking in another direction of the demonstration, they will not grasp it. This is why eye-tracking is useful.

Eye-tracking systems have flourished in the past few years due to their ease of use, high sensitivity, and especially their non-intrusiveness. It is a very useful tool in many research domains, including visualization, activity recognition, affect detection, and especially learning, to understand how students evolve and progress while learning. In fact, how students are reasoning while solving problems is a fundamental issue, regarding the technical and methodological challenges, due to the complexity of the reasoning process. Research has demonstrated that a relationship exists between *eye movements* and *cognitive processes* (Hunter et al., 2010). As a result, the use of eye-tracking can be beneficial as a means to identify the link between the task-relevant information and the learner's knowledge state (BenKheder et al., 2019).

Combining EEG and eye-tracking through neurofeedback methods should be a challenge for future methods of learning assessment. Neurofeedback can be used to enable people to learn to manage specific aspects of their neuronal activity and develop skills for self-regulation of brain activity, through a brain–computer interface.

In the same way that the evolution of theories has gradually converged towards the individual, the present and the future of the evaluation of learning will pass by measurements of the cerebral activities and eye-tracking which are more and more refined thanks to technological means under development. A vast field of research is opening on new learning strategies.

11. INTELLIGENT MENTORING SYSTEMS: TAPPING INTO AI TO DELIVER THE NEXT GENERATION OF DIGITAL LEARNING

Vania Dimitrova

In the past decades, the successful and deployable solutions of intelligent learning environments have predominantly been in formal education, usually school or university settings. In the next decades, I believe that AI in Education (AIED) will be established strongly in the broader educational landscape, with regard to a range of learning contexts, such as informal learning, adult education, professional learning, to name but a few. I argue that this calls for a change in the way we envisage the role of AIED and the way we design AIED systems. I believe that this will lead to a new family of AIED systems that provide support for coaching and mentoring.

Changing the Learning Landscape

Changes in a field, which can lead to new paradigms, often come from external drivers linked to socioeconomic and technological developments. A key driver calling for a change in our view of intelligent learning environments is the new learning landscape. There are strong voices for education to prepare for the rapidly changing skills demand. For example, the World Economic Forum lists key skills needed for jobs, such as analytical thinking and innovation, active learning and learning strategies, complex problem solving, critical thinking, and creativity. The Council of Europe's key competences for living in democratic societies and participating in democratic culture include: values (e.g. valuing human dignity and human rights, cultural diversity, democracy, and fairness), attitudes (e.g. openness to cultural otherness and to other beliefs, respect, civic-mindedness and self-efficacy), skills (e.g. autonomous learning skills, listening and observing, empathy, flexibility and adaptability, cooperation, and conflict-resolution), knowledge and critical understanding (e.g. critical understanding of the self, of language and communication, and of the world, human rights, culture, environment, and sustainability). This broadens the education scope to address soft (transferable) skills and to include lifelong learning in real contexts and everyday life experiences. At the same time, there is a growing demand for re-skilling, as stated by the futurist Alvin Toffler: "The illiterate of the 21st century will not be those who cannot read and write, but those who cannot learn, unlearn, and relearn." Hence, stronger emphasis is being put on lifelong learning, self-regulation, and developing growth mindset. Professional learning models aimed at developing reflexive practitioners who can make meaning of their practical experience and become lifelong learners, are becoming endorsed in workplace-based learning.

Mentoring

Mentoring is seen as a highly effective method to support the development of transferable skills, to increase motivation and confidence, and to develop self-regulation and self-determination. Because of insufficient space, I will not review the various definitions of mentoring and coaching and the nuances of differences between them. Hereinafter, I will use the term *mentor* as a generic term for the various interactions one can have with a more experienced person or with peers helping us to make meaning from our experience, develop self-awareness, become more engaged and motivated, set goals and reflect on their achievement, develop confidence and self-esteem, and thus realise one's full potential. Mentoring support can be offered by tutors, supervisors, friendsand particularly dedicated mentors and coaches. Mentoring can take different forms: 1:1 (the traditional model of mentor and mentee), 1:M (one mentor for a group of mentees), M:M (peer mentoring, usually in collaborative settings), M:1 (experience shared by many people in helping a mentee). Technology can play a key role by providing scalable and cost-effective support, tailored to the different mentoring models and adapted to individual needs and contexts.

Diversity-Aware Intelligent Mentoring Systems

Today, a vast number of learners use technology for educational purposes; and this will apply for future uses of technology to support mentoring. Consequently, technology support for mentoring would need to consider learner diversity, including demographic diversity (gender, age, culture), functional diversity (different experiences, different knowledge), and subjective diversity (personality characteristics, traits, values, motivations, beliefs). Educators and technologies supporting the education process would need to think about a wider scope of learner characteristics beyond knowledge and experience. A key challenge then would be how to understand diversity at scale. Digital transformation of education enables the capturing of digital traces which can be analysed to understand the learners, their behaviour, and the overall learning process. Hence, learning analytics approaches can provide "sensors" about aspects of the learners and the learning context.

This can then indicate situations in which we can provide nudges to foster behaviour change and facilitate learning. Nudges are widely used in behaviour change interventions (e.g. health, driving, green living) to influence choice without restricting any options. In mentoring systems, nudges can direct the learner to aspects that are beneficial for learning (e.g. to promote reflection, to improve engagement), yet preserve the freedom of exploration and the notion of self-control. Learning analytics combined with nudges can augment the mentoring models.

1:1 mentoring

We can provide intelligent personal mentors that act as conversational agents capable of "sensing" relevant situations about the learner (e.g. disengagement, lack of confidence, lack of motivation, boredom, procrastination, frustration) and offer interactive nudges (e.g. prompts, hints, questions, tips, suggestions) to help the learner become self-aware of their behaviour and identify what and how they may change. Another direction for using intelligent personal mentors is learning from experience – an effective approach for soft skills learning and for professional learning. Through the new technologies, such as speech, sensing devices, information about real-world experience can be captured. The personal mentor can then detect relevant situations in the digital traces with learners' experience and help them set goals and devise ways to achieve these goals.

1:M mentoring

These are group mentoring contexts done by one mentor, which refers to group-based interactive systems or collaborative spaces. Intelligent support can be offered to both the mentor and the mentees. We can envisage "sensors", in the form of group behaviour analytics, to understand individual differences and group dynamics. This can then inform nudges to be sent to individual members or to the group as a whole. Nudges can be offered automatically to the learners or can be suggested to a human mentor to help them moderate the group discussions.

M:M mentoring

Consider diversity-aware social interaction spaces where learners share experiences, knowledge, tips. As with 1:M, we can envisage "sensors" of group behaviour to detect interesting situations and nudges to offer direction that can change behaviour to benefit learning. Different contexts – soft skills learning, professional learning, informal learning – can utilise this approach. One of the main challenges in group contexts with diverse learners (e.g. cross-generational, cultural differences) is to ensure that the virtual social spaces are inclusive, that is, that everyone can benefit from their social experience. Analysis of the group interaction data can identify critical moments where members are ignored, may be confused, or become disengaged. The system can interfere to make the collective experience beneficial for the individual (e.g. raise group awareness, foster links, promote knowledge exchange of knowledge and experience).

M:1 mentoring

This refers to learning from collective intelligence. We can envisage nudging for informal learning where the learner freely explores a large information space with contributions from other people (e.g. comments, tips, stories, shared resources). In this case, domain models (ontologies) can help to map content to concepts from the domain, in order to design navigation nudges that direct the learner (offering signposts or recommendations) to content that can expand their knowledge and help them learn from others' experience. Learning through information exploration is being explored by the information retrieval community. In the future, closer links with the AIED community will enable more robust models that turn exploratory search tasks into effective informal learning scenarios.

Research Directions

To realise the intelligent mentoring system vision, we will need interdisciplinary partnerships that link researchers from computer scientists, professional learning, cognitive science, social science, and management. Possible research questions that these partnerships can consider include:

- **Learner needs:** What are the key lifelong learning challenges which require mentoring? What digital technologies do learners and tutors use to tackle these challenges? What are the main barriers (e.g. ethics, security, reliability)?
- **Technology:** What are the potential platforms/technologies that can be utilised for mentoring? What are the strengths and limitations of these platforms? What mentor-like features can we add to existing systems and how?
- **Learner/group sensing:** What individual differences should be considered; and how can they be captured/utilised in a holistic contextual model? What digital traces about the learner experience can be collected? What intelligent data analytics methods can be used to derive high-level behaviour models from low-level interaction data?
- **Domain modelling:** What "critical" situations do human mentors recognise when advising mentees? What contextual factors do human mentors take into account? What "nudging interventions" do human mentors use?
- **Intelligent nudges:** Can we shape a generic nudging framework (e.g. aspects which may include "sensors" to detect problem behaviour, interaction and navigation nudges to foster target behaviour)? What individual differences shall we consider for effective nudges?

- **Validation:** How reliable are the computational models? What are the pedagogical benefits of intelligent mentors, for example, evidence of reflection and self-awareness, improved confidence and self-esteem, better self-regulated learning, improved motivation? Who benefits and who does not?
- **Impact:** What findings can inform adopting mentor-like features in intelligent systems? How to turn proof-of-concept prototypes into deployable intelligent mentoring systems?

NOTE

1. For the purposes of this chapter, we consider the default role for a PA to be as a tutor, coach, or teacher. In other situations, such as role playing or peer learning, other factors may play a more important role in their design (e.g., in language training, agents would be designed to accurately represent a foreign country).

REFERENCES

Adib, A., Norman, A., & Orji, R. (2021). Persuasive application for discouraging unhealthy gaming behaviour. In *2021 IEEE 9th International Conference on Serious Games and Applications for Health (SeGAH)* (pp. 1–9). https://ieeexplore.ieee.org/abstract/document/9551891/.

Adorni, G., Alzetta, C., Koceva, F., Passalacqua, S., & Torre, I. (2019). Towards the identification of propaedeutic relations in textbooks. In *International Conference on Artificial Intelligence in Education* (pp. 1–13).

Agrawal, S., & Goyal, N. (2012, June). Analysis of Thompson sampling for the multi-armed bandit problem. In *Conference on Learning Theory* (p. 39). JMLR Workshop and Conference Proceedings.

Ahn, J. (2002). Beyond single equation regression analysis: Path analysis and multi-stage regression analysis. *The American Journal of Pharmaceutical Education*, 66(1), 37–41.

Aleven, V., Sewall, J., Andres, J. M., Sottilare, R., Long, R., & Baker, R. (2018, June). Towards adapting to learners at scale: Integrating MOOC and intelligent tutoring frameworks. In *Proceedings of the Fifth Annual ACM Conference on Learning at Scale* (pp. 1–4).

Allard, D., & Mizoguchi, R. (2021). Dr. Mosaik: A holistic framework for understanding the English tense-aspect system based on ontology engineering. *Research and Practice in Technology Enhanced Learning*, 16(1), 1–38.

Alpizar-Chacon, I., & Sosnovsky, S. (2021). Knowledge models from PDF textbooks. In *New Review of Hypermedia and Multimedia* (pp. 1–49).

Azevedo, R., Taub, M., & Mudrick, N. V. (2017). Understanding and reasoning about real-time cognitive, affective, and metacognitive processes to foster self-regulation with advanced learning technologies. In *Handbook of Self-Regulation of Learning and Performance* (pp. 254–270). Routledge.

Baker, R. S., Ogan, A. E., Madaio, M., & Walker, E. (2019). Culture in computer-based learning systems: Challenges and opportunities. *Computer-Based Learning in Context*, 1(1), 1–13. https://doi.org/10.5281/zenodo.4057223.

Barria-Pineda, J., Brusilovsky, P., & He, D. (2019). Reading mirror: Social navigation and social comparison for electronic textbooks. In *First Workshop on Intelligent Textbooks at 20th International Conference on Artificial Intelligence in Education* (pp. 30–37). http://ceur-ws.org/Vol-2384/paper03.pdf.

Basu, S., Biswas, G., Sengupta, P., Dickes, A., Kinnebrew, J. S., & Clark, D. (2016). Identifying middle school students' challenges in computational thinking-based science learning. *Research and Practice in Technology Enhanced Learning*, 11(1), 1–35.

Baylor, A. L. (2011). The design of motivational agents and avatars. *Educational Technology Research and Development*, 59(2), 291–300. https://doi.org/10.1007/s11423-011-9196-3.

Baylor, A. L., & Kim, Y. (2004). Pedagogical agent design: The impact of agent realism, gender, ethnicity, and instructional role. In J. C. Lester, R. M. Vicari, & F. Paraguaçu (Eds.), *Intelligent Tutoring Systems*. Heidelberg.

Ben Khedher, A., Jraidi, I., & Frasson, C. (2019). Predicting learners' performance using EEG and eye tracking features. In *The 32nd International FLAIRS Conference*, May 19–22, 2019, Florida, USA.
Berka, C., Levendowski, D. J., Lumicao, M. N., et al. (2007). EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation, Space, and Environmental Medicine,* 78, B231–B244.
Biswas, G., Segedy, J. R., & Bunchongchit, K. (2016). From design to implementation to practice – A learning by teaching system: Betty's brain. *International Journal of Artificial Intelligence in Education*, 26(1), 350–364.
Blanchard, E., & Mizoguchi, R. (2014). Designing culturally-aware tutoring systems with MAUOC, the more advanced upper ontology of culture. *Research and Practice in Technology Enhanced Learning*, 9(1), 41–69.
Blanchard, E. G. (2012, June). On the WEIRD nature of ITS/AIED conferences. In *International Conference on Intelligent Tutoring Systems* (pp. 280–285). Berlin, Heidelberg: Springer.
Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How People Learn* (Vol. 11). National Academy Press.
Brusilovsky, P., Eklund, J., & Schwarz, E. (1998). Web-based education for all: A tool for developing adaptive courseware. In H. Ashman & P. Thistewaite (Eds.), *Seventh International World Wide Web Conference* (pp. 291–300).
Bulger, M. (2016). Personalized learning: The conversations we're not having. *Data and Society,* 22(1), 1–29.
CATS. (2015). Sixth international workshop on culturally-aware tutoring systems. https://archium.ateneo.edu/discs-faculty-pubs/175/.
Chaouachi, M., & Frasson, C. (2010). Exploring the relationship between learners EEG mental engagement and affect, short paper. In *ITS 2010: 10th International Conference on Intelligent Tutoring Systems*. Pittsburgh, Pennsylvania: Springer Verlag, June 14–18, 2010.
Chaouachi, M., Jraidi, I., & Frasson, C. (2011). Modeling mental workload using EEG features for intelligent systems. In *User Modeling, Adaptation and Personalization, UMAP 2011*, Girona, Spain, July 2011.
Chaudhri, V. K., Cheng, B., Overtholtzer, A., Roschelle, J., Spaulding, A., Clark, P., Greaves, M., & Gunning, D. (2013). Inquire biology: A textbook that answers questions. *AI Magazine*, 34(3), 55–72.
Cian, H., Dou, R., Castro, S., Palma-D'souza, E., & Martinez, A. (2021). Facilitating marginalized youths' identification with STEM through everyday science talk: The critical role of parental caregivers. *Science Education*. Advance online publication. https://doi.org/10.1002/sce.21688.
Craig, S. D., & Schroeder, N. L. (2018). Design principles for virtual humans in educational technology environments. In *Deep Comprehension* (pp. 128–139). Routledge.
Dill-Shackleford, K. E., Ramasubramanian, S., Behm-Morawitz, E., Scharrer, E., Burgess, M. C., & Lemish, D. (2017). Social group stories in the media and child development. *Pediatrics*, 140(Supplement 2), S157–S161.
Dimitrova, V., & Mitrovic, A. (2021). Choice architecture for nudges to support constructive learning in active video watching. *International Journal of Artificial Intelligence in Education*. Advance online publication. https://doi.org/10.1007/s40593-021-00263-1.
D'Mello, S. K. (2016). Giving eyesight to the blind: Towards attention-aware AIED. *International Journal of Artificial Intelligence in Education*, 26(2), 645–659.
Dolog, P., Henze, N., Nejdl, W., & Sintek, M. (2004). The personal reader: Personalizing and enriching learning resources using semantic web technologies. In P. M. E. Bra & W. Nejdl (Eds.), *Third International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems* (pp. 85–94). Lecture Notes in Computer Science. Springer-Verlag.
Dou, R., Hazari, Z., Dabney, K., Sonnert, G., & Sadler, P. (2019). Early informal STEM experiences and STEM identity: The importance of talking science. *Science Education*, 103(3), 623–637.
Ebben, M., & Murphy, J. S. (2014). Unpacking MOOC scholarly discourse: A review of nascent MOOC scholarship. *Learning, Media and Technology*, 39(3), 328–345.
Emdin, C., Adjapong, E., & Levy, I. (2016). Hip-hop based interventions as pedagogy/therapy in STEM: A model from urban science education. *Journal for Multicultural Education*, 10(3), 307–321.
Ferrini-Mundy, J. (2013). Driven by diversity. *Science*, 340(6130), 278.

Finkelstein, S., Yarzebinski, E., Vaughn, C., Ogan, A., & Cassell, J. (2013). The effects of culturally congruent educational technologies on student achievement. *International Conference on Artificial Intelligence in Education*.
Frasson, C., & Heraz, A. (2011). Emotional learning. In *Encyclopedia of the Sciences of Learning*. Springer Verlag.
Gallagher, M. (2019). Artificial intelligence and the mobilities of inclusion: The accumulated advantages of 5G networks and surfacing outliers. In *Artificial Intelligence and Inclusive Education* (pp. 179–194). Springer.
Garner, P. W., Gabitova, N., Gupta, A., & Wood, T. (2018). Innovations in science education: Infusing social emotional principles into early STEM learning. *Cultural Studies of Science Education*, 13(4), 889–903.
Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. In Wheeler (Ed.), *Review of Personality and Social Psychology* (Vol. 1, pp. 141–165). Sage.
Goleman, D. (Ed.). (2006). *Emotional Intelligence, Why It Can Matter More Than IQ*. Bentam. ISBN-13 9780553804911.
Gynther, K. (2016). Design framework for an adaptive MOOC enhanced by blended learning: Supplementary training and personalized learning for teacher professional development. *Electronic Journal of e-Learning*, 14(1), 15–30.
Hachey, A. C. (2020). Success for all: Fostering early childhood STEM identity. *Journal of Research in Innovative Teaching & Learning*, 13(1), 135–139. https://doi.org/10.1108/JRIT-01-2020-0001.
Heffernan, N. T., & Heffernan, C. L. (2014). The ASSISTments ecosystem: Building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. *International Journal of Artificial Intelligence in Education*, 24(4), 470–497.
Heidig, S., & Clarebout, G. (2011). Do pedagogical agents make a difference to student motivation and learning? *Educational Research Review*, 6(1), 27–54.
Heraz, A., & Frasson, C. (2009). Predicting learner answers correctness through brainwaves assessment and emotional dimensions. In *AIED'2009: 14th International Conference on Artificial Intelligence in Education*. IOS Press.
Hofstede, G. (2022). Hofstede's cultural dimensions theory. https://en.wikipedia.org/wiki/Hofstede%27s_cultural_dimensions_theory.
Hohl, H., Böcker, H.-D., & Gunzenhäuser, R. (1996). Hypadapter: An adaptive hypertext system for exploratory learning and programming. *User Modeling and User-Adapted Interaction*, 6(2–3), 131–156.
Holstein, K., & Doroudi, S. (2021). Equity and artificial intelligence in education: Will "AIEd" amplify or alleviate inequities in education? *arXiv preprint arXiv:2104.12920*.
Hunter, M., Mach, Q. H., & Grewal, R. S. (2010). The relationship between scan path direction and cognitive processing. In *Proceedings of the Third C* Conference on Computer Science and Software Engineering* (pp. 97–100).
Hutchins, N. M., Biswas, G., Maróti, M., Lédeczi, Á., Grover, S., Wolf, R., … McElhaney, K. (2020a). C2STEM: A system for synergistic learning of physics and computational thinking. *Journal of Science Education and Technology*, 29(1), 83–100.
Hutchins, N. M., Biswas, G., Zhang, N., Snyder, C., Lédeczi, Á., & Maróti, M. (2020b). Domain-specific modeling languages in computer-based learning environments: A systematic approach to support science learning through computational modeling. *International Journal of Artificial Intelligence in Education*, 30, 537–580.
Hutchins, N. M., Snyder, C., Emara, M., Grover, S., & Biswas, G. (2021). Analyzing debugging processes during collaborative, computational modeling in science. In C. Hmelo-Silver, B. de Wever, & J. Oshima (Eds.), *14th International Conference on Computer-Supported Collaborative Learning – CSCL 2021* (pp. 221–224). International Society of the Learning Sciences.
Jackson, G. T., Boonthum, C., & McNamara, D. S. (2010, June). The efficacy of iSTART extended practice: Low ability students catch up. In *International Conference on Intelligent Tutoring Systems* (pp. 349–351). Berlin, Heidelberg: Springer.
Johnson, W. L., & Lester, J. C. (2018). Pedagogical agents: Back to the future. *AI Magazine*, 39(2), 33–44.
Kavcic, A. (2004). Fuzzy user modeling for adaptation in educational hypermedia. *IEEE Transactions on Systems, Man, and Cybernetics*, 34(4), 439–449.

Kinnebrew, J., Loretz, K., & Biswas, G. (2013). A contextualized, differential sequence mining method to derive students' learning behavior patterns. *Journal of Educational Data Mining*, 5(1), 190–219.

Kinnebrew, J., Segedy, J., & Biswas, G. (2017). Integrating model-driven and data-driven techniques for analyzing learning behaviors in open-ended learning environments. *IEEE Transactions on Learning & Technology*, 10(2), 140–153.

Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling*. Guilford Publications.

Kulik, C. L. C., & Kulik, J. A. (1982). Effects of ability grouping on secondary school students: A meta-analysis of evaluation findings. *American Educational Research Journal*, 19(3), 415–428.

Kumar, G., Banchs, R. E., & D'Haro, L. F. (2015). Revup: Automatic gap-fill question generation from educational texts. In *Tenth Workshop on Innovative Use of NLP for Building Educational Applications* (pp. 154–161).

Kumpulainen, K., & Wray, D. (Eds.). (2003). *Classroom Interaction and Social Learning: Frome Theory to Practice*. Routledge.

Lane, H. C., Cahill, C., Foutz, S., Auerbach, D., Noren, D., Lussenhop, C., & Swartout, W. (2013). The effects of a pedagogical agent for informal science education on learner behaviors and self-efficacy. In K. Yacef, H. C. Lane, & J. P. Mostow (Eds.), *Proceedings of the 16th International Conference on Artificial Intelligence in Education (AIED2013)* (Vol. 6738, pp. 309–318). Springer. https://doi.org/10.1007/978-3-642-39112-5_32.

Lane, H. C., & Schroeder, N. L. (2022). Pedagogical agents. In B. Lugrin, C. Pelachaud, & D. Traum (Eds.), *The Handbook on Socially Interactive Agents: 20 Years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics* (Vol. 2). ACM Press.

Leelawong, K., & Biswas, G. (2008). Designing learning by teaching agents: The Betty's brain system. *International Journal of Artificial Intelligence in Education*, 18(3), 181–208.

Leinhardt, G., Crowley, K., & Knutson, K. (2003). *Learning Conversations in Museums*. Taylor & Francis.

Liu, H., Ma, W., Yang, Y., & Carbonell, J. (2016). Learning concept graphs from online educational data. *Journal of Artificial Intelligence Research*, 55, 1059–1090.

Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology*, 106(4), 901.

MacLeod, L., Suruliraj, B., Gall, D., Bessenyei, K., Hamm, S., Romkey, I., Bagnell, A., Mattheisen, M., Muthukumaraswamy, V., Orji, R., & Meier, S. (2021). A mobile sensing app to monitor youth mental health: Observational pilot study. *JMIR mHealth and uHealth,* 9(10), e20638.

McClure, E. R., Guernsey, L., Clements, D. H., Bales, S. N., Nichols, J., Kendall-Taylor, N., & Levine, M. H. (2017). STEM starts early: Grounding science, technology, engineering, and math education in early childhood. In *The Joan Ganz Cooney Center at Sesame Workshop.*

McElhaney, K. W., Zhang, N., Basu, S., McBride, E., Biswas, G., & Chiu, J. L. (2020). Using computational modeling to integrate science and engineering curricular activities. In *Proceedings of the International Conference of the Learning Sciences (ICLS*) (Vol. 3, pp. 1357–1364). Nashville, TN, USA.

McLaren, B. M., DeLeeuw, K. E., & Mayer, R. E. (2011). Polite web-based intelligent tutors: Can they improve learning in classrooms? *Computers & Education*, 56(3), 574–584.

McReynolds, A. A., Naderzad, S. P., Goswami, M., & Mostow, J. (2020). Toward learning at scale in developing countries: Lessons from the global learning XPRIZE field study. In *Proceedings of L@S '20: Seventh ACM Conference on Learning @Scale*, August 12–14, 2020, Virtual Event, USA.

Melis, E., Andres, E., Büdenbender, J., Frischauf, A., Goguadze, G., Libbrecht, P., & Ullrich, C. (2001). ActiveMath: A generic and adaptive web-based learning environment. *International Journal of Artificial Intelligence in Education*, 12(4), 385–407.

Mizoguchi, R. (2020). A proposal for a new framework of learner modeling: From modeling "understanding" to "not understanding". *The Journal of Information and Systems in Education*, 19(1), 9–14.

Mizoguchi, R., & Bourdeau, J. (2016). Using ontological engineering to overcome AI-ED problems: Contribution, impact, and perspectives. *International Journal of Artificial Intelligence in Education*, 26(1), 91–106.

Moreno, K. N., Person, N. K., Adcock, A. B., Eck, R., Jackson, G. T., & Marineau, J. C. (2002). Etiquette and efficacy in animated pedagogical agents: The role of stereotypes. In *AAAI Symposium on Personalized Agents*, Cape Cod, MA.

Moreno, R., Mayer, R., & Lester, J. (2000). Life-like pedagogical agents in constructivist multimedia environments: Cognitive consequences of their interaction. In *World Conference on Educational Multimedia, Hypermedia and Telecommunications.*

Mostow, J. (2019). RoboTutor. www.robotutor.org.

Munshi, A., Rajendran, R., Ocumpaugh, J., Biswas, G., Baker, R. S., & Paquette, L. (2018). Modeling learners cognitive and affective states to scaffold SRL in open ended learning environments. In *Proceedings of the International Conference on User Modelling, Adaptation and Personalization* (UMAP) (pp. 131–138).

Nkambou, R., Mizoguchi, R., & Bourdeau, J. (2010). *Advances in Intelligent Tutoring Systems.* Springer-Verlag.

NSF INCLUDES Coordination Hub. (2020). Broadening participation in STEM: Evidence-based strategies for improving equity and inclusion of individuals in underrepresented racial and ethnic groups. (Research Brief No. 1).

Nye, B. D. (2014, June). Barriers to ITS adoption: A systematic mapping study. In *International Conference on Intelligent Tutoring Systems* (pp. 583–590). Cham: Springer.

Nye, B. D. (2015). Intelligent tutoring systems by and for the developing world: A review of trends and approaches for educational technology in a global context. *International Journal of Artificial Intelligence in Education*, 25(2), 177–203.

Ogan, A., Yarzebinski, E., Fernández, P., & Casas, I. (2015, June). Cognitive tutor use in Chile: Understanding classroom and lab culture. In *International Conference on Artificial Intelligence in Education* (pp. 318–327). Cham: Springer.

Orji, F. A., & Vassileva, J. (2021). Modelling and quantifying learner motivation for adaptive systems: Current insight and future perspectives. In *International Conference on Human-Computer Interaction* (pp. 79–92).

Orji, F. A., Vassileva, J., & Greer, J. (2021). Evaluating a persuasive intervention for engagement in a large university class. *IJAIED*, 27(4), 700–725. https://doi.org/10.1007/s40593-021-00260-4.

Ostrow, K., Heffernan, N., & Williams, J. J. (2017). Tomorrow's EdTech today: Establishing a learning platform as a collaborative research tool for sound science. *Teachers College Record*, 119(3), 1–36.

Oyibo, K., Orji, R., & Vassileva, J. (2017). The influence of culture in the effect of age and gender on social influence in persuasive technology. In *Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization* (pp. 47–52). https://dl.acm.org/doi/abs/10.1145/3099023.3099071.

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, 8, 422.

Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2008). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest*, 9(3), 105–119.

Patikorn, T., & Heffernan, N. T. (2020, August). Effectiveness of crowd-sourcing on-demand assistance from teachers in online learning platforms. In *Proceedings of the Seventh ACM Conference on Learning@ Scale* (pp. 115–124).

Pearl, J., & Mackenzie, D. (2018). *The Book of Why: The New Science of Cause and Effect*. Basic Books.

Picard, W. R. (Eds.). (1997). *Affective Computing.* MIT Press.

Prihar, E., Patikorn, T., Botelho, A., Sales, A., & Heffernan, N. (2021, June). Toward personalizing students' education with crowdsourced tutoring. In *Proceedings of the Eighth ACM Conference on Learning@ Scale* (pp. 37–45).

Rajendran, R., Kumar, A., Carter, K. E., Levin, D. T., & Biswas, G. (2018). Predicting learning by analyzing eye-gaze data of reading behavior. *International Educational Data Mining Society* (pp. 455–461).

Razzaq, L. M., & Heffernan, N. T. (2009, June). To tutor or not to tutor: That is the question. In *Proceedings of the 2009 Conference on Artificial Intelligence in Education: Building Learning Systems That Care: From Knowledge Representation to Affective Modelling* (pp. 457–464). Amsterdam: IOS Press.

Razzaq, R., Ostrow, K. S., & Heffernan, N. T. (2020, July). Effect of immediate feedback on math achievement at the high school level. In *International Conference on Artificial Intelligence in Education* (pp. 263–267). Cham: Springer.

Reeves, B., & Nass, C. (1996). *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge University Press.

Renninger, K. A., Bachrach, J. E., & Hidi, S. E. (2019). Triggering and maintaining interest in early phases of interest development. *Learning, Culture and Social Interaction*, 23, 100260.

Renninger, K. A., & Hidi, S. (2016). *The Power of Interest for Motivation and Engagement*. Routledge.

RoboTutor LLC. (2019). RoboTutor_2019. https://github.com/RoboTutorLLC/RoboTutor_2019.

Rodrigo, M. M. T. (2021, July). Impediments to AIS adoption in the Philippines. In *International Conference on Human-Computer Interaction* (pp. 412–421). Cham: Springer.

Rodrigues, L., Palomino, P. T., Toda, A. M., Klock, A. C. T., Oliveira W., Avila-Santos, A. P., Gasparini, I., & Isotani, S. (2021). Personalization improves gamification: Evidence from a mixed-methods study. *Proceedings of the ACM Human-Computer Interaction*, 5(CHI PLAY), Article 287. https://dl.acm.org/doi/abs/10.1145/3474714.

Roschelle, J., Feng, M., Murphy, R. F., & Mason, C. A. (2016). Online mathematics homework increases student achievement. *AERA Open,* 2(4), 2332858416673968.

Savard, I., & Mizoguchi, R. (2019). Context or culture: What is the difference? *Research and Practice in Technology Enhanced Learning*, 14(1), 1–12.

Schroeder, N. L., & Adesope, O. O. (2014). A systematic review of pedagogical agents' persona, motivation, and cognitive load implications for learners. *Journal of Research on Technology in Education*, 46(3), 229–251.

Schroeder, N. L., & Adesope, O. O. (2015). Impacts of pedagogical agent gender in an accessible learning environment. *Journal of Educational Technology & Society*, 18(4), 401–411.

Schroeder, N. L., Adesope, O. O., & Gilbert, R. B. (2013). How effective are pedagogical agents for learning? A meta-analytic review. *Journal of Educational Computing Research*, 49(1), 1–39.

Segedy, J. R., Kinnebrew, J. S., & Biswas, G. (2015). Using coherence analysis to characterize self-regulated learning behaviors in open-ended learning environments. *Journal of Learning Analytics*, 2, 13–48.

Snyder, C., Hutchins, N., Biswas, G., Emara, M., Grover, S., & Conlin, L. (2019). Analyzing students' synergistic learning processes in physics and CT by collaborative discourse analysis. In *Proceedings of the International Conference on Computer Supported Collaborative Learning* (pp. 360–367). Lyon, France.

Sociology Group. (2022). What is culture: Basic elements of culture and features. https://www.sociologygroup.com/elements-of-culture-basic-elements-of-culture/.

Sosnovsky, S. (2014). Math-bridge: Closing gaps in European remedial mathematics with technology-enhanced learning. In *Mit werkzeugen mathematik Und stochastik lernen – Using Tools for Learning Mathematics and Statistics* (pp. 437–451). Springer Spektrum.

Sosnovsky, S., Hsiao, I. H., & Brusilovsky, P. (2012, September). Adaptation "in the wild": ontology-based personalization of open-corpus learning material. In *European Conference on Technology Enhanced Learning* (pp. 425–431). Berlin, Heidelberg: Springer.

Steenbergen-Hu, S., & Cooper, H. (2014). A meta-analysis of the effectiveness of intelligent tutoring systems on college students' academic learning. *Journal of Educational Psychology,* 106(2), 331.

Taub, M., Azevedo, R., Rajendran, R., Cloude, E. B., Biswas, G., & Price, M. J. (2021). How are students' emotions related to the accuracy of cognitive and metacognitive processes during learning with an intelligent tutoring system? *Learning and Instruction*, 72, 101200.

Thaker, K., Zhang, L., He, D., & Brusilovsky, P. (2020). Recommending remedial readings using student's knowledge state. In *13th International Conference on Educational Data Mining*, July 10–13, 2020, pp. 233–244.

UNESCO. (2021). Leading SDG 4 – Education 2030. UNESCO Website. https://en.unesco.org/themes/education2030-sdg4.

Uy-Tioco, C. S. (2019). 'Good enough' access: Digital inclusion, social stratification, and the reinforcement of class in the Philippines. *Communication Research and Practice*, 5(2), 156–171.

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4), 197–221.

Veletsianos, G., & Shepherdson, P. (2016). A systematic analysis and synthesis of the empirical MOOC literature published in 2013–2015. *The International Review of Research in Open and Distributed Learning*, 17(2), 192–221.

Wang, S., Ororbia, A., Wu, Z., Williams, K., Liang, C., Pursel, B., & Giles, C. L. (2016). Using prerequisites to extract concept maps from textbooks. In *25th International Conference on Information and Knowledge Management* (pp. 317–326).

Weber, G., & Brusilovsky, P. (2001). ELM-ART: An adaptive versatile system for web-based instruction. *International Journal of Artificial Intelligence in Education*, 12(4), 351–384.

Winchell, A., Mozer, M., Lan, A., Grimaldi, P., & Pashler, H. (2018). Can textbook annotations serve as an early predictor of student learning? In *11th International Conference on Educational Data Mining*, Buffalo, USA, pp. 431–437.

XPRIZE. (2019). Empowering children to take control of their own learning. learning.xprize.org.

Yarbro, J. T., & Olney, A. M. (2021). Contextual definition generation. In *3rd Workshop on Intelligent Textbooks at 22nd International Conference on Artificial Intelligence in Education* (pp. 74–83). http://ceur-ws.org/Vol-2895/paper09.pdf.

Yin, C., Yamada, M., Oi, M., Shimada, A., Okubo, F., Kojima, K., & Ogata, H. (2019). Exploring the relationships between reading behavior patterns and learning outcomes based on log data from E-books: A human factor approach. *International Journal of Human–Computer Interaction*, 35 (4–5), 313–322.

Zhai, C., & Massung, S. (2016). *Text Data Management and Analysis: A Practical Introduction to Information Retrieval and Text Mining*. Morgan & Claypool.

Zhang, N., Biswas, G., Chiu, J. L., & McElhaney, K. W. (2019). Analyzing students' design solutions in an NGSS-aligned earth sciences curriculum. In *Proceedings of the 20th International Conference on Artificial Intelligence in Education (AIED)*, Chicago, pp. 532–543.

Zhang, N., Biswas, G., & Hutchins, N. M. (2021). Measuring and analyzing students' strategic learning behaviors in open-ended learning environments. *International Journal of Artificial Intelligence in Education*. Advance online publication. https://doi.org/10.1007/s40593-021-00275-x.