Transforming Educational Technology Through Convergence

Conference Report

Ryan S. Baker, Yeeva Cheng, Lizzie Jones, Seiyon Lee, Jaclyn Ocumpaugh, Malia Sieve, Ulrich Boser

15 December 2022

Summary

This year's release of NAEP scores revealed a large impact on student learning from COVID: the largest drops in reading and math in three decades of administering the tests. Even before the pandemic, NAEP scores were lagging. To get America's educational system back on track, we brought in 40 experts from a variety of groups — from educational technology companies to philanthropic organizations to teachers — to discuss potential solutions.

The group underscored the multidisciplinary and convergent nature of education, a field that cuts across psychology, cognitive science, sociology, and economics as well as the specific domains being learned (math, biology, chemistry, etc). Education has been traditionally siloed, often resistant to crucial societal innovations from technology to the changing nature of careers and work. This makes education an excellent potential track for a Convergence Accelerator, which "builds upon basic research and discovery to accelerate solutions toward societal impact."

After ideating deliverables in the key areas of data science education, middle school math, and assessment, the group discussed cross-cutting trends among them. They found it critically important to support convergence in education that will help prepare today's students to become informed decision-makers, engaged problem-solvers, and self-directed lifelong learners. This report surfaces the key themes and necessary partnerships that experts believe are critical for improvements to educational opportunities. It then examines the key disciplines and convergence required to produce deliverables capable of transforming the educational landscape in the US.

Key future directions for deliverables, their intellectual merits, and broader societal impacts:

● The middle school math deliverables focus on increasing student motivation, the relevance of math concepts and skills, supporting collaborative and project-based learning, optimizing and expanding feedback mechanisms, and developing AI to respond to students' inputs. These innovations will help unveil more about achievement and opportunity gaps and other mechanisms that affect groups of students differentially in STEM domains.

● The data science education deliverables focus on preparing students with procedural skills to work with data and supporting teachers to provide timely feedback on data science-related assessments. Intellectual merits of these deliverables include understanding the paths to how data science education will be integrated into mainstream curricula–or developed and taught as its own subject (Engel, 2017)–given its interdisciplinary nature.

● The assessment deliverables focused on the development of new, increasingly unobtrusive ways to assess students, including elements such as gamification and assessing a broader range of skills (such as self-regulation and collaborative learning). Intellectual merits of these deliverables include deeper understanding of learning processes, creating broader impacts through more valid, less disruptive, and more comprehensive assessments.

Background

The pandemic had a massive impact on student learning. The most recent NAEP scores, for instance, show a large drop in outcomes (National Center for Education Statistics, 2022a). In fact, it was the biggest drop in Math and English scores in over three decades of administering the tests.

While the pandemic led to a renewed focus on using learning technology, innovation in this area has yet to close the gap in student needs. Even though some exemplary, highly-effective educational technology platforms experienced up to tenfold increases in userbases (Gillespie et al., 2022; Baker et al., 2022), the pivot to emergency online education mostly consisted of attempting to teach in traditional ways but through Zoom (Bonk, 2020; McArthur, 2021), using credit recovery platforms and learning platforms already under fire for low quality (Ball & Grimaldi, 2021), or using untested approaches (Teräs et al., 2020; Aguilar, 2020).

These results underscore the need for transformative solutions in education. As noted in previous research, intractable long-standing problems often are intractable for good reason (National Research Council, 2014; Roco, 2016). They connect to broader societal problems, and their solutions require new thinking and new approaches. This type of new thinking can only be achieved by bringing together a combination of expertise, both technical and social: the type of combination that goes beyond simply a mixture of experts to collaborations that fully integrate different types of expertise. Simply put, transformation requires disciplinary convergence (National Research Council, 2014; Roco, 2016) is needed.

One potential for change in education comes from artificial intelligence (AI). AI has sparked impressive developments including intelligent tutoring systems, learning analytics dashboards, and dialogue tutors, sometimes realized through chatbots (Graesser et al., 2012; Verbert et al., 2013; Molenaar, 2022; Katz et al., 2021; Chen et al., 2022). Contemporary AI-based learning systems can detect and respond to students' knowledge, learning strategies, engagement, and emotion (Owen et al., 2019; Gowda et al., 2013; Bosch et al., 2016; D'Mello et al., 2017).

However, AI alone cannot solve educational inequities and can sometimes even perpetuate biases and inequities in education, due to algorithmic bias (Baker & Hawn, 2022). Careful efforts are needed to design effective and unbiased learning systems. Today, many of the most effective approaches to using AI in education are designed to have AI work in tandem with humans to respond to learners (Shum et al., 2021; Holstein & Aleven, 2021; Yang et al., 2021; Dimitriadis et al., 2021).

Increasingly, AI is used not only within learning platforms, but also to drive and support learning engineering efforts (Doignon & Falmagne, 2012; Rosé et al., 2019; Lomas et al., 2016), providing actionable information to learning designers and learning scientists (Holstein et al., 2019; Baker et al., 2018; Goodall & Kolodner, 2022). For example, data-driven interviewing is a form of convergent research where a qualitative educational researcher informs AI about uncommon but important classroom events, and then the AI informs the researcher in real-time that the event has just occurred (Ocumpaugh et al., 2021).

These new, convergent approaches show considerable potential for transforming education. But these approaches are also highly complex. Currently, it can be challenging to assemble the multidisciplinary research and development teams required to conduct the convergent work necessary for the success of these approaches. Thus, careful focus is needed on which contexts these methods will be applied to first. We propose three contexts where progress is feasible and needs are strong: mathematics education, data science education, and learning-integrated formative assessment. We address them in turn.

There has been concern about students' performance in mathematics in the United States for decades (National Commission on Excellence in Education, 1983). This concern has led to some effective curricula and technologies (Murphy et al., 2020; Pane et al., 2014; Donnelly & Parmar, 2022), but as a whole, mathematics performance in the United States has not improved as hoped. In fact, the pandemic-related drops in middle school math scores have raised concerns about learners' readiness for other academic topics that require foundational math skills (Kuhfeld et al., 2022). Research on intelligent tutoring systems, cognitive tutors, digital assistants, and other technologies in well-defined domains like math have offered some promising opportunities to differentiate or personalize learning, leading to improved outcomes (Koedinger and Corbett, 2006; Ritter et al., 2007; Canfield, 2001; Melis & Siekmann, 2014).

While there has been concern about mathematics education for decades, interest in data science education has been a relatively more recent development, matching the increased societal use of data (Wise, 2020; Lee & Wilkerson, 2021; Wilkerson & Polman, 2020). There is rapidly growing demand for data scientists and our daily lives are increasingly influenced by how data is used (Horton et al., 2015; Provost & Fawcett, 2013). Curricular support for data science has lagged behind the growing need, and many curricula still teach data analysis (if at all) using outdated representations such as stem-and-leaf plots. Students already interact with data on a daily basis, may have or are developing preconceived notions about data, and increasingly consume digital information (National Center for Education Research, 2021). Thus, there is a need for innovative learning platforms and approaches that can effectively engage learners in data science in order to prepare them for future careers and to thrive in the data revolution.

Across domains, supporting teachers in supporting students requires a step forward in assessment. Teachers cannot accurately gauge how their students are performing without good information. However, most teachers are still limited to information either from informal assessments (i.e., homework and tests) that they develop and grade themselves (Shute & Ventura, 2013), or to information from large-scale assessment systems that require pulling a student out of instruction for several class days, taking away essential instructional time while often assessing knowledge at a fairly shallow level using traditional items (Shute & Rahimi, 2017). A new generation of assessment systems that assess while students learn and tap deeper levels of understanding can inform teachers while immediately benefiting students.

In this report, we discuss the findings of an interdisciplinary conference of experts on potential directions for a new NSF Convergence Accelerator track that can speed progress in these areas and, ultimately, in education as a whole.

Convergence Research

Within this project, we organized a conference of 40 experts, spanning a range of disciplinary expertise and current professional roles (see Participant Profiles in Appendix A). In order to bring in the best possible collection of experts (many of whom are highly busy in their ongoing work), we structured this conference as a set of three spaced-out virtual sessions, prompting our experts to converge ideas and share expertise from their respective fields. These sessions began with an initial meeting that included all participants. Workshop organizers established a common understanding and emphasized the goals of the series. Participants were then divided into three tracks (middle school math, data science education, and assessments) with participants further divided into small groups that were diverse in background and expertise. For example, a team might include one member working in education, another in research, and another in philanthropy.

Subgroups were then tasked to coordinate among themselves to meet and ideate deliverables that could provide novel solutions to the most pressing issues in the subject area within their track's mission. Many of our groups chose to go beyond this initial request, articulating how common problems existed in more than one track. Subgroups were encouraged to especially focus on issues of educational access and equity, as well as steps developers would need to take to ensure that a deliverable did not introduce or reinforce bias. A guide (see Meeting 2 Guide in Appendix B) with prompting questions was provided to subgroups to capture other important reflections, such as the disciplines and types of expertise that would be needed for the development of the solution. Such questions prompted groups to emphasize the roles, needs, and limitations of stakeholders within the educational ecosystem, and in some cases, such considerations helped inform the product design. For example, a few teams emphasized the importance of ensuring that the deliverables designed were as unobtrusive in data collection processes as possible or effective in reducing teachers' grading workloads to enable them to focus on differentiating feedback.

During the third session, subgroups met with each other and gave each other critical feedback and recommendations, which is incorporated within our discussion of the themes and deliverables below. Additional experts who could not participate in the second meeting also joined these sessions, providing a greater range of expertise.

Themes

After observing the conference process and key deliverables, we identified cross-cutting key themes in the advancements in education that currently have traction and would benefit from further investment and support.

Adaptive learning systems: Adaptive — or personalized — approaches to learning, teaching, and assessment came up in many discussions. Intelligent tutoring systems are not a new technology; proven systems like MATHia and ALEKS have demonstrated high effectiveness (Craig et al., 2013), but a new generation of AI (Fernandez et al., 2022; Scarlatos et al., 2022) promises a new generation of adaptive learning that better adapts to individual differences and responds in richer, more flexible ways.

Specific deliverables included:

- Personalized learning experiences that use content that better motivates individual students to avoid confusing disengagement with lack of understanding
- Adaptive learning activities that not only identify weaknesses but also identify strengths and help students build on them
- Bringing adaptivity to a wider range of activities, including project-based, community-based, problem-based, game-based, and discovery-based learning

Technology-enhanced human tutoring: Expert human tutoring has long been seen as a highly-effective approach (VanLehn, 2011; Bloom, 1984), but it has been hard to scale (Khachatryan, 2020). Artificial intelligence methods can better inform tutors who may encounter a student only briefly, and may help select the best tutor out of a pool of tutors, for a specific student within a specific need. Helping human tutors to specialize on specific content or specific types of student needs can also lead to tutors achieving expert-level performance in a shorter time.

Specific deliverables included:

● Integrated systems to recommend and deploy mathematics tutoring resources within districts
● Informational support for human tutors and designs that help tutors

Generative AI for classroom use: Too much of education remains decontextualized, and students cannot always see the value of what they learn or the relevance to their lives (Duffy & Jonassen, 2013; Merill, 2013; Grabinger & Dunlap, 1995; Choi & Hanafin, 1995). The generative nature of next-generation AI tools — such as foundation models for language (Devlin et al., 2018; Floridi et al., 2020) —- can make it possible to generate learning experiences customized to a broader range of topics and student interests. These technologies have clear potential risks in terms of bias and fairness (Mehrabi et al., 2021; Brown et al., 2020) but also may be able to reduce bias by recognizing a broader range of student responses than earlier technologies could.

learning experiences customized to a broader range of topics and student interests. These technologies have clear potential risks in terms of bias and fairness (Mehrabi et al., 2021; Brown et al., 2020) but also may be able to reduce bias by recognizing a broader range of student responses than earlier technologies could.

Specific deliverables included:

● AI-based foundation models that can generate text can adapt content or assessment to an individual student's interests, and abstract back to the original concepts
● Technologies based on foundation models can be used to grade student papers on a range of topics and give higher-quality automated feedback
● Technologies based on foundation models can be used to create dialogue tutoring systems that can respond flexibly to a broader range of student inputs

Dashboards for classroom support: Dashboards and other data-interface tools can support teachers and students (and other school personnel) with clear and actionable information (Verbert et al., 2013; 2020; Sedrakyan et al., 2020). Their design can be supported with new convergent developments in data analytics, user engagement, and behavioral science. Such dashboards would rely on new assessments that are adaptive as well as provide learning experiences.

Specific deliverables included:

● New technologies that can provide teachers with actionable information they can use and make sense of
● Dashboards that would rely on assessments where students learn as they are assessed to create faster, more effective feedback loops

Collaborative and social learning: In many contexts, collaborative and social learning can engage students and help them learn to consider other perspectives on a topic (Dillenbourg et al., 1996; Pea, 1994; Koschmann, 2002; Scardamalia & Bereiter, 2006; Stahl, 2006). However, collaborative learning can be difficult to orchestrate in a classroom without carefully designed support (Dillenbourg et al., 2009; Krejins et al., 2003). Collaboration provides opportunities for articulation and elaboration, which are particularly important when problem-solving in STEM and data science.

Specific deliverables included:

● Games and other learning systems that facilitate collaboration
● Dashboards (and other reporting tools) that facilitate classroom orchestration efforts by teachers

Next-generation research methods:
Educational research and development have historically been fairly slow and inefficient. Improved infrastructure and more convergent methods can lead to faster discovery about what is and isn't working, and in turn can lead to faster improvements.

Specific deliverables included:

● A/B testing platforms like E-TRIALS that can support researchers from different disciplinary backgrounds (and teachers) in trying out new ideas and seeing if they work
● Public data sets can bring in a wider range of computer science researchers to improve the AI underpinning educational technologies. Properly cleaned and annotated with meta-data, they can also become useful tools for data science education.
● Greater interoperability between platforms that can make it possible to use information across platforms and offer dashboards with broader range of information to teachers
● Convergent research methods like data-driven classroom interviewing that can bring new types of researchers into the process of improving educational technologies

Equity: Inequity is linked to differences in opportunity, and subsequently, achievement. Given enduring equity issues, there was concern about next-generation educational technologies and how they might address such issues. Unfortunately, some high-tech solutions have the potential to reinforce inequity in several fashions (Madaio et al., 2022; Holmes et al., 2022; Holstein & Doroudi, 2019). Participants noted that careful attention needs to be paid to ensuring that new algorithms are not algorithmically biased; it is likely that there are biases that are currently not well-documented.

Specific deliverables included:

● Culturally-responsive content embedded within learning systems, and improved teacher training related to these issues

● Games and other learning systems with representation of multiple groups
● Research methods that deliberately seek to identify and reduce the risk of algorithmic biases
● The development of technologies that are easier for low-income students and families to access, including technologies that work on low-priced devices such as Chromebooks and phones —or even offline, for students who do not have home access to stable internet connections

Disciplines Required for Transformative Convergence

Given the focus of the convergence accelerator – and the highly multidisciplinary nature of the project – we outline here some of the different disciplines and communities that would be central to improving student outcomes via an education track.

Experts in teaching and learning: Experts in teaching and learning understand key considerations like how to scaffold educational content, which can help learners pace their consumption of content and assess their skills at appropriate levels of rigor. They also know how to connect new learning materials to the broader curriculum and how to engage diverse students. Thus, finding ways to bring in experts in the practical and everyday aspects of education — teachers, curriculum developers, and instructional designers, as well as school and district level administrators — will improve the resulting tools in ways that boost student outcomes.

Teachers along with school and district administrators are also critical to the uptake and scaling of pedagogical ideas into practice — if they are dissatisfied or uninterested in a new idea or technology, it will not get used. Their input on available resources will help inform key features of deliverables and how these best and most effectively respond to students' needs.

Technologists and experts in AI, ML, and NLP: While technology has had a major impact on society in recent years, it has yet to have a tremendous impact on student outcomes, outside of a small number of cases. But there is a lot of potential. The advances of technology could offer educators new ways to differentiate instruction and rapidly assess and intervene in students' learning at a faster pace. From intelligent tutoring systems to internet scaling, AI technology could potentially help augment and expand learning environments with different ways of assessing students, providing feedback, and personalizing instruction. Of course,technology does not come without potential dangers. Privacy and security experts, for instance, need to be at the forefront of questions around how to anonymize student data and which metadata should be collected and how it should be used.

Gaming industry: Direct instruction–the mode of education in which teachers stand at the front of a classroom and lecture–is increasingly known to be non-optimal, as research demonstrates the importance of hands-on activities and real-life scenarios. Learning games offer exciting ways to potentially reimagine both learning activities and assessment. Game-based assessments and activities, when integrated with findings and methods from learning science research, can produce learning experiences that can increase motivation, relevance, and even belonging for students. Game-based learning environments also have the potential to create opportunities for learners to exhibit and practice important self-regulation skills (i.e., planning and strategizing). Through using concepts from games and gamification and observing the behavioral, social, and emotional elements of learning, researchers and educators would be able to build a path toward lifelong learning.

Community groups: Education does not stop at the school. The ecosystem of effective education includes a wide array of community organizations from libraries to recreation centers to after-school clubs. Encouraging interaction between these organizations and those who are designing new learning systems could improve opportunities for learning substantially.

For example, many of the services traditionally offered by these programs extend the classroom (e.g., after-school tutoring programs). Making sure that these community organizations are connected with the most innovative learning systems that the AI community has developed could help ensure equitable access to motivating learning experiences (e.g., innovative learning games).

At the same time, the infrastructure that these organizations provide can create opportunities for researchers to collaborate with local communities. These collaborations are important for facilitating equitable designs. For example, the creation of appropriate avatars —and even unbiased analytics—requires working with diverse groups of learners through the design phases of online learning systems.

Data science community: Data science is a discipline essential for students to learn so that future generations maintain ownership of which, how, and for whom data is used. Looking forward, data science experts working on real-world problems could collaborate more deeply with educators and school systems to help students learn data science skills, from coding and analyzing data to critiquing the presentation and use of data in their everyday lives. There are already examples of this, including an initiative by the multinational software corporations SAP and Chevron and the San Francisco 49ers football team to help students learn data science (49ers.com, 2022).

Research communities: The explosion of educational tools and technologies, while offering options, also overwhelm educators with options. The process of "trying out" a new technology or tool without clear evidence of improved outcomes is often a point of frustration and lost resources. Research should continue to provide clarity around what works and what is worth the investment for educators and administrators. Additionally, the research community's engagement with educational technologies support rapid experimentation in which interventions that do not work are quickly ruled out while those with potential are increasingly refined and adapted, informing theory and practice. A third way that research might fuel convergence is its potential to inform programmatic or technological design features. During our workshop, groups often listed multidisciplinary combinations of learning science, computer science, and psychology research as disciplines that informed their recommendations about which deliverables were most likely to successfully blend the social with the technical to achieve positive change.

Partnerships

Partnerships will be key to the success of future learning technologies and approaches. We attempted to include representatives of key stakeholder groups in our workshop, and their participation demonstrated the high value of their inclusion:

• *Education industry.* Educational technology companies and nonprofit organizations were represented in the workshops, including ASSISTments, Khan Academy, CueThink, Bootstrap, Edtech Recharge, and Cambium Assessment. These companies represented a span of specializations ranging from math assessment platforms to apps focusing on improving critical thinking skills.

Industry will be central to a future convergence accelerator track because they can help engage users as well as scale successful approaches. Beyond the small set of workshop participants, key stakeholders will include education-specific organizations like the College Board but also larger corporations, such as Google and Microsoft, with broader goals that are active in classrooms. Bringing in these organizations will facilitate the scaled delivery of practical solutions that raise student outcomes. Organizations like gaming companies can support the development of interventions, while data science companies can help to outline future careers and the needs for learning technologies to support them.

• *Educator groups.* Within the workshops, educators were represented from multiple sectors of practice and hailing from diverse districts around the country, ranging from middle school math teachers to PK-12 instructional coordinators. Educators were critical to the dialogue around the needs of diverse learners that should influence innovation trends, the barriers to introducing technologies and tools into the classroom setting, and the challenges of collecting data. Educators emphasized the fact that schools and districts are inundated with educational products, materials, and programs, but they are unsure of the research, theory and evidence backing decision paths to product uptake.
A potential track would go much further, of course, and educator groups at multiple levels would need to be included in decisions and research on the programs, interventions, and technologies that are integrated into school settings. At the grassroots level, teachers are directly impacted by these decisions that can help or hinder their efforts. Their perspectives provide insights to what would economize time, how to provide them information in actionable and useful ways, and what interventions are feasible. Administrators at the school and district levels would also need to be part of the work. Their knowledge of standards alignment, assessment procedures and policies, and local resources would help facilitate the design of deliverables that will be used and the programmatic management of interventions.

• *Philanthropic support.* Many participants represented private philanthropies including the Walton Family Foundation, the Bill & Melinda Gates Foundation, Valhalla Foundation, and Schmidt Futures. Philanthropic organizations could potentially play a key role in a track, helping to support the work. Many of these private foundations have supported work in these areas in the past and could share key "lessons learned." These organizations can also fund projects that the NSF might not typically fund, such as networks that help build the learning engineering ecosystem overall.

• *Research community.* The largest proportion of participants included researchers with specific projects with expertise varying from automated writing evaluation to game-based learning. Research community members were helpful in thinking through key ideas such as data collection and predictive modeling Offering deep and diverse expertise, the research community would continue to play a driving role in educational advancements in a potential convergence track. For example, there is high potential to create more cross-cutting research between experts in data science, learning science, psychology, and pedagogy. Key stakeholders for this change would involve researchers in universities, nonprofits, and private industry working together and with other types of stakeholders. Though universities have essential expertise, it is worth going beyond the academic community with insights from applied researchers in industry who work more directly with technologists and other stakeholders experienced in scaling and troubleshooting a product or deliverable.

Deliverables

Given the focus of the convergence accelerator – and the highly multidisciplinary nature of the project – we outlined some of the different disciplines and communities that would be central to improving student outcomes via an education track.

Middle school math: Within middle school math, there was a broad range of deliverables proposed, from increasing students' motivation through clarifying the relevance of math concepts and skills, to supporting collaborative learning, to bringing adaptivity to a broader range of activities (i.e. project-based, community-based, problem-based, game-based, discovery-based), to optimizing and expanding feedback mechanisms. A number of proposed deliverables were noteworthy in terms of feasibility, scalability, and potential impact.

For instance, one group discussed the value of creating a dataset of handwritten math problem solutions. Currently, there are substantial amounts of valuable data on learning and problem-solving processes that are available in students' handwriting, but this data is mostly used for grading by teachers. Hand grading is time-consuming, delays feedback, and makes it difficult to see important trends across a student's assignments (which the teacher may grade weeks apart) (Moon et al., 2022). The group suggested creating a public dataset of tagged images and labels that would facilitate developing technology to recognize students' handwriting and automate feedback, reducing teachers' workloads tremendously while leveraging convergent trends in computer vision, math pedagogy, and human-computer interaction.

Another team argued for a new convergent approach to designing computer-based intelligent tutoring systems, going past the current generation of technologies that typically get one type of adaptivity right (see Baker, 2016). This "super tutor" would combine the strengths of different approaches, supporting not just learning of material but also strengthening motivation/attitudes towards math and problem-solving through selection of material and narrative-based activities. It could include a human-in-the-loop and collaborative design elements, whereby a teacher could view a dashboard of students clustered by common need. This would allow teachers to assign student groups that could more independently problem-solve while focusing time on subsets of students who would benefit most from personalized attention. Such a system could go beyond identifying students' weaknesses, to also recognizing students' strengths and helping build on them.

A third deliverable discussed technologies and programs that support students finding mathematics in the world around them. A high-tech approach might involve augmented reality glasses (or phone-based augmented reality like Pokemon Go) that help students answer questions like "how tall is that building?" or "how many blueberries are there in that container?" Such an AI system could recognize the opportunity for instruction (with general detectors of arrays of objects, for example), understand mathematical activities based on this recognition (such as counting and multiplying to estimate the number of blueberries or using trigonometry to answer the height question), and scaffold the student (and perhaps their parents as well) in solving the problem. Cognizant that a high-tech approach might not reach all students, a low-tech variation of this deliverable could involve new approaches to co-design that help teachers understand how to guide students in real-world problem-solving.

Data science education: A range of deliverables were proposed in the data science track to give students improved data science skills. Across groups, there was a focus on increasing student motivation and clarifying the relevance of data science topics to students' lives. Many groups emphasized the potential of using AI to build curricular resources that assist teachers in providing timely feedback as they build students' data science skills in core areas such as data cleaning and analysis.

A key current challenge to data science education is the lack of structured datasets containing the organized metadata that facilitate data science practice (Kross & Guo, 2019; Bart et al., 2018). While there are open data sites, many of these include datasets that require intensive cleaning or are noisy in ways that are unpredictable (Bart et al., 2018; Finzer, 2013), presenting challenges to novice learners who do not yet have data cleaning skills. To solve this issue, one deliverable was to create a greater number of pre-packaged datasets — still based on authentic data — that are already cleaned or present noise in predictable ways, and where the context of the data is clearly communicated, allowing students to focus more on analysis. One group suggested that these cleaned datasets could be made available for use by middle school students by modifying an existing platform for data science competitions (Anslow et al., 2016; Finzer, 2013), a type of "Kaggle for Kids," to include both publicly available cleaned datasets and notebooks of code for users to apply data science skills and conduct analysis.

Other groups raised the broader challenge of the lack of integration of data science education into current schooling. Discussion noted that, while many educators are eager to teach about data science, there is a shortage of K-12 educators with the knowledge or expertise to teach foundational topics in data science (Weiland & Engledowl, 2022). One proposed deliverable would address this limitation by creating an AI-assisted platform that would allow teachers to search and filter by building up queries such as, "I want to teach regression to 9th graders using examples from sports."

The tool could conceivably generate new modules by combining the desired method with a relevant dataset or finding video resources from a database. This would enable the teacher to focus on working with their students rather than customizing or writing new code, which is often a time-intensive barrier. Through focusing on themes and topics that are most interesting to students, the deliverable could increase the motivation and relevance of data science education to students, while allowing an entry point for more novice teachers to introduce data science education or integrate it easily within their subject area. Other deliverables involved using intelligent tutor technologies or chatbots to help teachers quickly develop their own mastery of data science prior to teaching it to students (perhaps just-in-time for specific topics, the day before a teacher teaches a specific topic).

Groups also discussed the need to teach cross-cutting themes in artificial intelligence, as well as methods; discussion was particularly focused on the recent and growing evidence for algorithmic bias (Aket et al., 2021; Buolamwini & Gebru, 2018; Noble, 2018). Bringing together a convergent community of experts to design standards for what should be covered in these areas will be important to develop students who can not only use data science methods but also think critically about them.

Assessment: Within groups discussing the future of assessment in education, there was tremendous excitement for new approaches that could improve the assessment data available to teachers, conduct valid assessment in the context of richer and more authentic activities, while simultaneously promoting learning rather than separating assessment from learning.

This discussion led to several notable deliverables. For example, one group discussed the use of foundation models for language, such as BERT and GPT-3 (Devlin et al., 2018; Floridi & Chiriatti, 2020), to assess open-ended student responses. Math assessments frequently include open-ended questions, such as “Write a rule that explains your results” or “Explain your answer,” that ask students to demonstrate their knowledge. The use of these types of prompts aligns with research suggesting that these types of items help unearth misconceptions and are effective for developing students’ competency in math concepts, but evaluating open-ended responses costs teachers immense amounts of time (Hancock, 1995). An assessment application could use a foundation model to identify patterns in students’ responses and identify frequently occurring misconceptions at both the individual and class levels, providing immediate feedback to students while supporting teachers in refining the instruction they offer. This type of technology could be applied both in classrooms and on summative assessments, guiding policy and enriching the breadth of understanding about students at a national level.

Another idea proposed the use of AI methods for automated item generation. Developing and validating assessment items can be time-intensive (Bechard et al., 2019), especially if attention is paid to developing items that are culturally relevant, unbiased, and engaging for students. Generative AI methods can help develop a range of high-quality, culturally responsive assessment items that range in cognitive complexity. This would allow students more dynamic assessment experiences where they could demonstrate learning in different ways, authentic and relevant to a diverse range of learners. This would provide educators a richer set of information than

students. Generative AI methods can help develop a range of high-quality, culturally responsive assessment items that range in cognitive complexity. This would allow students more dynamic assessment experiences where they could demonstrate learning in different ways, authentic and relevant to a diverse range of learners. This would provide educators a richer set of information than is currently possible. Furthermore, a hybrid human-computer system could be developed to leverage what individual students find motivating or engaging, using intelligent recommender system technology. Students could select what topics they find most applicable, providing greater agency to students, and the technology could use psychometric methods to ensure equivalence between different items. Alternatively, affect and engagement detection technology (D’Mello et al., 2017) could be used to identify when an item is functioning poorly for specific learners or specific groups of learners.

A fourth idea that emerged in the sessions was the use of modern AI approaches to assess a broader range of constructs, providing data to teachers and school leaders that goes beyond simply assessing academic competencies. Group work is a key component of authentic learning and assessment, for instance, but it is often challenging for teachers to support, document, and assess group work in real-time.

Moreover, several groups emphasized the importance of collaborative learning in providing equitable educational opportunities. Underrepresented minority students especially see themselves less represented in STEM and benefit from collaborations that facilitate belonging and peer support (Hatfield et al, 2022; Kricorian et al., 2020). Promoting collaborative learning in a way that is measurable not only furthers the research on best practice in orchestrating collaboration, but also supports diverse learners to co-regulate in ways that may feel more authentic to their contexts and approaches to problem-solving (Perry et al, 2017; Kricorian et al., 2020).

The components are in place for the convergence of disciplines in studying and scaffolding collaborative learning, including multi-dimensional, theoretical models of group dynamics that support or hinder collaborative learning; automated tools such as eye-tracking, affective and attentional computing, speech recognition to document group documents; and analytic methods such as nonlinear time series analysis, discourse modeling, and machine learning to leverage data.

Recent convergent projects have also demonstrated the feasibility of detecting students’ learning strategies (Hutt et al., 2021; Zhang et al., 2022; Azevedo et al., 2011; Azevedo & Witherspoon, 2008), and their current capacities and strategies for self-regulated learning and emotional self-regulation (Sabourin et al., 2013). Assessing this could provide valuable information for teachers and could support the development of a new generation of intelligent learning technologies that scaffold students in developing these skills. These technologies could assess (and therefore support) key 21st-century skills such as collaboration and creativity (Shute & Rahimi, 2021; Kim & Shute, 2015).

Conclusion/Discussion

The opportunity to bring together 40 experts resulted in fruitful discussion of the needs, trends, emerging technologies, and future directions for improving three key areas in education: middle school math, data science education, assessment. This dialogue underscored the need for convergence research, which we summarize below, providing an overview of future directions for deliverables, their intellectual merit, and their broader impacts.

Middle School Math

● **Future directions of deliverables:** Future educational approaches and technologies based on the convergence of disciplinary areas have the potential to prepare middle schoolers with foundational math skills as well as addressing the social and environmental factors that have particularly affected underrepresented minority students' persistence in math education. Specifically, there are highly promising trends in increasing student motivation and the relevance of math concepts and skills, supporting collaborative and project-based learning, optimizing and expanding feedback mechanisms, and developing AI that can respond to a wider range of types of student input.

● **Intellectual merit:** While experts across the different fields represented in this report know about the achievement and opportunity gaps that differentially affect groups of students in STEM domains, the causal mechanisms behind the persistence of these problems require a multidisciplinary approach to address, particularly as new algorithmic inequities emerge. The deliverables outlined above include technologies with the built-in infrastructure to not only understand these mechanisms, but to respond in real-time to these issues, inequities, and challenges.

● **Broader impacts:** Through supporting middle school math education that is relevant to students' lives and integrates social learning, student motivation for math will increase. Optimizing and expanding feedback mechanisms will help focus teachers' efforts, enabling them to more easily differentiate instruction and focus their time and attention on students who most benefit from it.

Data Science Education

● **Future directions of deliverables:** The convergence of different areas and communities has the potential to substantially improve how we teach data science while increasing motivation and the relevance of data science topics to students' lives. It may be possible to accelerate the uptake of top-quality data science education by developing curricular resources that assist teachers in providing timely feedback as they build students' procedural skills in data cleaning and analysis.

● **Intellectual merit:** Data science education is needed to prepare students to become informed, data literate citizens of the future. A simultaneous barrier and opportunity within data science is how much it integrates computational skills, social sciences, and mathematical concepts. The paths to how data science education is integrated into school core subjects–or if it will someday become its own class entirely–are yet to be determined. Investing in research and deliverables mentioned above extends the key literature, pedagogical practices, and policy directions for data science education at an essential time.

● **Broader impacts:** Through supporting data science education that emphasizes increasing motivation and relevance and using AI to build resources and assist teachers, we can build data science curricula that teach this complex domain drawing upon many disciplines and skill sets. Data science is an inherently convergent field, making it hard for many teachers and students; a convergent approach to teaching it can prepare students for careers and lives where understanding data is paramount.

Assessment

● **Future directions of deliverables:** The potential impact of convergent research in the assessment is likely to be highly transformative. Improving opportunities for convergence across different fields of research has the potential to improve the validity of future assessments, and will facilitate the development of assessments that could be used in new (and increasingly unobtrusive) ways. These include the development of assessments in new contexts (including games) that assess a broader range of skills such as self-regulation and collaborative learning.

● **Intellectual merits:** Learning engineering and learning analytics can support a new-generation of assessment that produces rich data at fine-grained levels to reveal more about learning processes. The convergence between psychology, psychometrics, and data science promise a future of assessment that is more valid, less disruptive, and more comprehensive.

● **Broader impacts:** Through continued research and development, the next generation of assessments can more reliably measure what students know, how they learn, and how instruction can be adjusted in terms of cognitive, metacognitive, self-regulatory and socio-emotional skills that students need to learn most effectively. These new assessments can also provide teachers with real-time information that is less disruptive to the learning process than traditional forms of assessment.

Acknowledgement

The authors acknowledge NSF #2231524 “Conference: Transforming Educational Technology Through Convergence” for financial support.

References

49ers.com (2021) 49ers EDU Partners with SAP to Add Data Science to Curriculum. Retrieved 12/7/2022 from https://www.49ers.com/news/49ers-edu-partners-with-sap-to-add-data-science-to-curric ulum

Aguilar, S. J. (2020). A research-based approach for evaluating resources for transitioning to teaching online. Information and Learning Sciences.

Akter, S., McCarthy, G., Sajib, S., Michael, K., Dwivedi, Y. K., D’Ambra, J., & Shen, K. N. (2021). Algorithmic bias in data-driven innovation in the age of AI. International Journal of Information Management, 60, 102387.

Anslow, C., Brosz, J., Maurer, F., & Boyes, M. (2016, February). Datathons: an experience report of data hackathons for data science education. In Proceedings of the 47th ACM Technical Symposium on Computing Science Education (pp. 615-620).

Azevedo, R., Bouchet, F., Harley, J. M., Feyzi-Behnagh, R., Trevors, G., Duffy, M., ... & Yang, W. (2011). MetaTutor: An Intelligent Multi-Agent Tutoring System Designed to Detect, Track, Model Foster Self-Regulated Learning. In Proceedings of the Fourth Workshop on Self-Regulated Learning in Educational Technologies.

Azevedo, R., & Witherspoon, A. M. (2008, January). Detecting, Tracking, and Modeling Self-Regulatory Processes during Complex Learning with Hypermedia. In AAAI Fall Symposium: Biologically Inspired Cognitive Architectures (pp. 16-26).

Baker, R. S., Gowda, S. M., & Salamin, E. (2018). Modeling the learning that takes place between online assessments. In Proceedings of the 26th international conference on computers in education (pp. 21-28).

Baker, R. S., Nasiar, N., Gong, W., & Porter, C. (2022). The impacts of learning analytics and A/B testing research: a case study in differential scientometrics. International Journal of STEM Education, 9(1), 1-10.

Ball, S. J., & Grimaldi, E. (2021). Neoliberal education and the neoliberal digital classroom. Learning, Media and Technology, 1-15.

Bart, A. C., Kafura, D., Shaffer, C. A., & Tilevich, E. (2018, February). Reconciling the promise and pragmatics of enhancing computing pedagogy with data science. In Proceedings of the 49th ACM Technical Symposium on Computer Science Education (pp. 1029-1034).

Bechard, S., Clark, A., Swinburne Romine, R., Karvonen, M., Kingston, N., & Erickson, K. (2019). Use of evidence-centered design to develop learning maps-based assessments. International Journal of Testing, 19(2), 188-205.

Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. Educational Researcher, 13(6), 4–16.

Bonk, C. J. (2020). Pandemic ponderings, 30 years to today: Synchronous signals, saviors, or survivors?. Distance Education, 41(4), 589-599.

Bosch, N., D’Mello, S. K., Baker, R. S., Ocumpaugh, J., Shute, V., Ventura, M., ... & Zhao, W. (2016, January). Detecting student emotions in computer-enabled classrooms. In IJCAI (pp. 4125-4129).

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. Advances in neural information processing systems, 33, 1877-1901.

Buckingham Shum, S., Ferguson, R., & Martinez-Maldonado, R. (2019). Human-centred learning analytics. Journal of Learning Analytics, 6(2), 1-9.

Buolamwini, J., & Gebru, T. (2018, January). Gender shades: Intersectional accuracy disparities in commercial gender classification. In Conference on fairness, accountability and transparency (pp. 77-91). PMLR.

Catalyzing a New Field: Data Science Education in K-12. (n.d.). 27.

Canfield, W. (2001). ALEKS: A Web-based intelligent tutoring system. Mathematics and Computer Education, 35(2), 152.

Chapman, B. (2022, October 24). Math Scores Dropped in Every State During Pandemic, Report Card Shows. Wall Street Journal. https://www.wsj.com/articles/math-scores-dropped-in-every-state-during-pandemic-report-card-shows-11666584062

Chapter 21. (n.d.). Society for Learning Analytics Research (SoLAR). Retrieved November 14, 2022, from https://www.solaresearch.org/publications/hla-17/hla17-chapter21/

Chen, X., Zou, D., Xie, H., Cheng, G., & Liu, C. (2022). Two Decades of Artificial Intelligence in Education. Educational Technology & Society, 25(1), 28-47.

Choi, J. I., & Hannafin, M. (1995). Situated cognition and learning environments: Roles, structures, and implications for design. Educational technology research and development, 43(2), 53-69.

Craig, S. D., Hu, X., Graesser, A. C., Bargagliotti, A. E., Sterbinsky, A., Cheney, K. R., & Okwumabua, T. (2013). The impact of a technology-based mathematics after-school program using ALEKS on student’s knowledge and behaviors. Computers & Education, 68, 495-504.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.

Diebold, G., & Han, C. (n.d.). How AI Can Improve K-12 Education in the United States. 33.

Dillenbourg, P., Baker, M., Blaye, A., & O'Malley, C. (1996). The evolution of research on collaborative learning. In E. Spada & P. Reiman (Eds.), Learning in Humans and Machine: Towards an interdisciplinary learning science (pp. 189–211). Oxford: Elsevier

Dillenbourg, P., Järvelä, S., & Fischer, F. (2009). The evolution of research on computer-supported collaborative learning. In Technology-enhanced learning (pp. 3-19). Springer, Dordrecht.

Dimitriadis, Y., Martínez-Maldonado, R., & Wiley, K. (2021). Human-centered design principles for actionable learning analytics. In Research on E-Learning and ICT in Education (pp. 277-296). Springer, Cham. diSessa, A. A. (2014). A history of conceptual change research: Threads and fault lines.

Doignon, J. P., & Falmagne, J. C. (2012). Knowledge spaces. Springer. Donnelly, S., & Parmar, R. (2022). A Comparison of IXL Math Online Software and Textbook Problems for Homework In An Urban Middle School. Journal of Computers in Mathematics and Science Teaching, 41(1), 45-67.

Dorn, E., Hancock, B., Sarakatsannis, J., & Viruleg, E. (2021). COVID-19 and education: The lingering effects of unfinished learning. McKinsey & Company, 27.

Duffys, T. M., & Jonassen, D. H. (2013). Constructivism: New implications for instructional technology. In Constructivism and the technology of instruction (pp. 1-16). Routledge.

Engel, J. (2017). Statistical literacy for active citizenship: A call for data science education. Statistics Education Research Journal, 16(1), 44-49.

Fernandez, N., Ghosh, A., Liu, N., Wang, Z., Choffin, B., Baraniuk, R., & Lan, A. (2022). Automated Scoring for Reading Comprehension via In-context BERT Tuning. arXiv preprint arXiv:2205.09864.

Finzer, W. (2013). The data science education dilemma. Technology Innovations in Statistics Education, 7(2).

Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. Minds and Machines, 30(4), 681-694.

Gillespie, J., Winn, K., Faber, M., & Hunt, J. (2022). Implementation of a Mathematics Formative Assessment Online Tool Before and During Remote Learning. In International Conference on Artificial Intelligence in Education (pp. 168-173). Springer, Cham.

Goodell, J., & Kolodner, J. (Eds.). (2022). Learning Engineering Toolkit: Evidence-Based Practices from the Learning Sciences, Instructional Design, and Beyond. Taylor & Francis.

Gowda, S. M., Baker, R. S., Corbett, A. T., & Rossi, L. M. (2013). Towards automatically detecting whether student learning is shallow. International journal of artificial intelligence in education, 23(1), 50-70.

Grabinger, R. S., & Dunlap, J. C. (1995). Rich environments for active learning: A definition. ALT-J, 3(2), 5-34.

Graesser, A. C., Conley, M. W., & Olney, A. (2012). Intelligent tutoring systems. APA educational psychology handbook, Vol 3: Application to learning and teaching., 451-473

Hancock, C. L. (1995). Implementing the assessment standards for school mathematics: Enhancing mathematics learning with open-ended questions. The Mathematics Teacher, 88(6), 496-499.

Hatfield, N., Brown, N., & Topaz, C. M. (2022). Do introductory courses disproportionately drive minoritized students out of STEM pathways? PNAS Nexus, 1(4), pgac167. https://doi.org/10.1093/pnasnexus/pgac167

Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., ... & Koedinger, K. R. (2022). Ethics of AI in education: Towards a community-wide framework. International Journal of Artificial Intelligence in Education, 32(3), 504-526.

Holstein, K., & Aleven, V. (2022). Designing for human–AI complementarity in K-12 education. AI Magazine, 43(2), 239-248.

Holstein, K., & Doroudi, S. (2019, March). Fairness and equity in learning analytics systems (FairLAK). In Companion proceedings of the ninth international learning analytics & knowledge conference (LAK 2019) (pp. 1-2).

Holstein, K., McLaren, B. M., & Aleven, V. (2019). Co-designing a real-time classroom orchestration tool to support teacher–AI complementarity. Journal of Learning Analytics, 6(2).

Horton, N. J. (2015). Challenges and opportunities for statistics and statistical education: looking back, looking forward. The American Statistician, 69(2), 138-145.

Hutt, S., Ocumpaugh, J., Ma, J., Andres, A. L., Bosch, N., Paquette, L., ... & Baker, R. S. (2021). Investigating SMART Models of Self-Regulation and Their Impact on Learning. International Educational Data Mining Society.

Khachatryan, G. A. (2020). Instruction modeling: Developing and implementing blended learning programs. Oxford University Press.

Katz, S., Albacete, P., Chounta, I. A., Jordan, P., McLaren, B. M., & Zapata-Rivera, D. (2021). Linking dialogue with student modelling to create an adaptive tutoring system for conceptual physics. International Journal of Artificial Intelligence in Education, 31(3), 397-445.

Kim, Y. J., & Shute, V. J. (2015). Opportunities and challenges in assessing and supporting creativity in video games. Video games and creativity, 99-117.

Kirschner, P. A., Sweller, J., Kirschner, F., & Zambrano R., J. (2018). From Cognitive Load Theory to Collaborative Cognitive Load Theory. International Journal of Computer-Supported Collaborative Learning, 13(2), 213–233. https://doi.org/10.1007/s11412-018-9277-y

Koschmann, T. (2002). Dewey's contribution to the foundations of CSCL research. In Proceedings of Computer Support for Collaborative Learning 2002 (pp. 17–22). Boulder, CO.

Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research. Computers in human behavior, 19(3), 335-353.

Kross, S., & Guo, P. J. (2019, May). Practitioners teaching data science in industry and academia: Expectations, workflows, and challenges. In Proceedings of the 2019 CHI conference on human factors in computing systems (pp. 1-14).

Kuhfeld, M., Soland, J., & Lewis, K. (2022). Test score patterns across three COVID-19-impacted school years. EdWorkingPaper: 22-521, 37-62.

Kuhfeld, M., Soland, J., & Lewis, K. (2022). Test Score Patterns Across Three COVID-19-Impacted School Years. Educational Researcher, 0013189X2211091. Lee, A. M. P. and N. T. (2019, September 26). AI is coming to schools, and if we're not careful, so will its biases. Brookings. https://www.brookings.edu/blog/the-avenue/2019/09/26/ai-is-coming-to-schools-and-i f-were-not-careful-so-will-its-biases/

Lee, V. R., Wilkerson, M. H., & Lanouette, K. (2021). A call for a humanistic stance toward K–12 data science education. Educational Researcher, 50(9), 664-672.

Lomas, J. D., Forlizzi, J., Poonwala, N., Patel, N., Shodhan, S., Patel, K., ... & Brunskill, E. (2016, May). Interface design optimization as a multi-armed bandit problem. In Proceedings of the 2016 CHI conference on human factors in computing systems (pp. 4142-4153).

Machine Learning and Higher Education. (n.d.). Retrieved October 19, 2022, from https://er.educause.edu/articles/2017/12/machine-learning-and-higher-education

Madaio, M., Blodgett, S. L., Mayfield, E., & Dixon-Román, E. (2022). Beyond "fairness": Structural (in) justice lenses on AI for education. In The Ethics of Artificial Intelligence in Education (pp. 203-239). Routledge.

McArthur, J. A. (2021). From classroom to Zoom room: Exploring instructor modifications of visual nonverbal behaviors in synchronous online classrooms. Communication Teacher, 1-12.

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. ACM Computing Surveys (CSUR), 54(6), 1-35.

Melis, E., & Siekmann, J. (2004, June). Activemath: An intelligent tutoring system for mathematics. In International Conference on Artificial Intelligence and Soft Computing (pp. 91-101). Springer, Berlin, Heidelberg.

Merrill, M. D. (2013). Constructivism and instructional design. In Constructivism and the technology of instruction (pp. 99-114). Routledge.

Miyake, N., & Kirschner, P. A. (2014). The Social and Interactive Dimensions of Collaborative Learning. In R. K. Sawyer (Ed.), The Cambridge Handbook of the Learning Sciences (2nd ed., pp. 418–438). Cambridge University Press.

Molenaar, I. (2022). Towards hybrid human-AI learning technologies. European Journal of Education, 57(4), 632-645.

Moons, F., Vandervieren, E., & Colpaert, J. (2022). Atomic, reusable feedback: a semi-automated solution for assessing handwritten tasks? A crossover experiment with mathematics teachers. Computers and Education Open, 3, 100086.

Murphy, R., Roschelle, J., Feng, M., & Mason, C. A. (2020). Investigating efficacy, moderators and mediators for an online mathematics homework intervention. Journal of Research on Educational Effectiveness, 13(2), 235-270.

National Center for Education Statistics (2022a) NAEP Long-Term Trend Assessment Results: Reading and Mathematics. Retrieved November 14, 2022, from https://www.nationsreportcard.gov/highlights/ltt/2022/

National Center for Education Statistics (2022b). NAEP Mathematics: National Achievement-Level Results. Retrieved October 28, 2022, from https://www.nationsreportcard.gov/mathematics/nation/achievement/?grade=8

National Research Council. (2014). Convergence: Facilitating transdisciplinary integration of life sciences, physical sciences, engineering, and beyond. National Academies Press.

Noble, S. U. (2018). Algorithms of oppression. In Algorithms of Oppression. New York University Press

Ocumpaugh, J., Hutt, S., Andres, J. M. A. L., Baker, R. S., Biswas, G., Bosch, N., ... & Munshi, A. (2021, November). Using Qualitative Data from Targeted Interviews to Inform Rapid AIED Development. In 29th International Conference on Computers in Education Conference, ICCE 2021 (pp. 69-74). Asia-Pacific Society for Computers in Education.

Osborne, J. (2007). Science Education for the Twenty First Century. 12.

Owen, V. E., Roy, M. H., Thai, K. P., Burnett, V., Jacobs, D., Keylor, E., & Baker, R. S. (2019). Detecting Wheel-Spinning and Productive Persistence in Educational Games. International educational data mining society.

Pane, J. F., Griffin, B. A., McCaffrey, D. F., & Karam, R. (2014). Effectiveness of cognitive tutor algebra I at scale. Educational Evaluation and Policy Analysis, 36(2), 127-144.

Pea, R. D. (1994). Seeing what we build together: Distributed multimedia learning environments for transformative communications. The Journal of the Learning Sciences, 3(3), 219–225.

Pellegrino, J. W. (2022). A Learning Sciences Perspective on the Design and Use of Assessment in Education. In R. K. Sawyer (Ed.), The Cambridge Handbook of the Learning Sciences (3rd ed., pp. 238–258). Cambridge University Press.

Pelletier, K., McCormack, M., Reeves, J., Robert, J., & Arbino, N. (2022). 2022 EDUCAUSE Horizon Report, Teaching and Learning Edition. 58.

Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. Big data, 1(1), 51-59.

Ritter, S., Anderson, J. R., Koedinger, K. R., & Corbett, A. (2007). Cognitive Tutor: Applied research in mathematics education. Psychonomic Bulletin & Review, 14(2), 249–255.

Roco, M. C. (2016). Principles and methods that facilitate convergence. In Handbook of science and technology convergence (pp. 17-41). Springer, Cham.

Rosé, C. P., McLaughlin, E. A., Liu, R., & Koedinger, K. R. (2019). Explanatory learner models: Why machine learning (alone) is not the answer. British Journal of Educational Technology, 50, 2943–2958.

Sabourin, J. L., Rowe, J. P., Mott, B. W., & Lester, J. C. (2013). Considering Alternate Futures to Classify Off-Task Behavior as Emotion Self-Regulation: A Supervised Learning Approach. Journal of Educational Data Mining, 5(1), 9-38.

Scardamalia, M., & Bereiter, C. (2006). Knowledge building. The Cambridge.

Scarlatos, A., Brinton, C., & Lan, A. (2022). Process-BERT: A Framework for Representation Learning on Educational Process Data. arXiv preprint arXiv:2204.13607.

Sedrakyan, G., Malmberg, J., Verbert, K., Järvelä, S., & Kirschner, P. A. (2020). Linking learning behavior analytics and learning science concepts: Designing a learning analytics dashboard for feedback to support learning regulation. Computers in Human Behavior, 107, 105512.

Shute, V. J., & Rahimi, S. (2017). Review of computer-based assessment for learning in elementary and secondary education. Journal of Computer Assisted Learning, 33(1), 1-19.

Shute, V. J., & Rahimi, S. (2021). Stealth assessment of creativity in a physics video game. Computers in Human Behavior, 116, 106647.

Shute, V. J. & Ventura, M. (2013). Measuring and supporting learning in games: Stealth assessment. Cambridge, MA: The MIT Press

Stahl, G. (2006). Group cognition: Computer support for building collaborative knowledge. Cambridge, MA: MIT Press. Teräs, M., Suoranta, J.,

Teräs, H., & Curcher, M. (2020). Post-Covid-19 education and education technology 'solutionism': A seller's market. Postdigital Science and Education, 2(3), 863-878.

Using Self-Regulated Learning as a Framework for Creating Inclusive Classrooms for Ethnically and Linguistically Diverse Learners in Canada | Request PDF. (n.d.). Retrieved November 14, 2022, from https://www.researchgate.net/publication/313472598_Using_Self-Regulated_Learning_as_a_Framework_for_Creating_Inclusive_Classrooms_for_Ethnically_and_Linguistically_ Diverse_Learners_in_Canada

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

Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos, J. L. (2013). Learning analytics dashboard applications. American Behavioral Scientist, 57(10), 1500-1509.

Verbert, K., Ochoa, X., De Croon, R., Dourado, R. A., & De Laet, T. (2020, March). Learning analytics dashboards: The past, the present and the future. In Proceedings of the tenth international conference on learning analytics & knowledge (pp. 35-40).

Weiland, T., & Engledowl, C. (2022). Transforming curriculum and building capacity in K–12 data science education. Harvard Data Science Review, 4(4).

Wilkerson, M. H., & Polman, J. L. (2020). Situating data science: Exploring how relationships to data shape learning. Journal of the Learning Sciences, 29(1), 1-10.

Wise, A. F. (2020). Educating data scientists and data literate citizens for a new generation of data. Journal of the Learning Sciences, 29(1), 165-181.

Wise, A. F., & Shaffer, D. W. (2015). Why Theory Matters More than Ever in the Age of Big Data.Journal of Learning Analytics, 2(2), 5–13.

Yang, S. J., Ogata, H., Matsui, T., & Chen, N. S. (2021). Human-centered artificial intelligence in education: Seeing the invisible through the visible.Computers and Education: Artificial Intelligence, 2, 100008.

Zhang, J., Andres, J. M., Alexandra, L., Hutt, S., Baker, R. S., Ocumpaugh, J., ... & Young, T. (2022). Detecting SMART Model Cognitive Operations in Mathematical Problem-Solving Process. InProceedings of the 15th InternationalConference on Educational Data Mining(p. 75).

Appendices

Appendix A: Participant Profiles

Transforming Educational Technology Through Convergence
Participant List

Assessments

First Name	Last Name	Title	Organization
Stephen	Aguilar	Assistant Professor of Education	USC Rossier School of Education
Jodi	Asbell-Clark	Director of EdGE	TERC
Scott	Crossley	Professor of Applied Linguistics and English as a Second Language	Georgia State University
Danielle	Eisenberg	Consultant	Walton Family Foundation
Britte	Haugan Cheng	Principal and Founder	Menlo Education Research
YJ	Kim	Assistant Professor of Design, Creative, and Informal Education	University of Wisconsin
Rene	Kizilcec	Assistant Professor of Information Science	Cornell University
Diane	Litman	Professor of Computer Science	University of Pittsburgh
Susan	Lottridge	Chief Scientist, Natural Language Acquisitions	Cambium Assessments
Danielle	McNamara	Professor, Psychology	Arizona State University
Xavier	Ochoa	Assistant Professor of Learning Analytics	NYU Steinhardt School of Culture, Education, and Human Development
Brandee	Tate	Program Officer, Assessment	Bill & Melinda Gates Foundation
John	Whitmer	Senior Fellow	Federation of American Scientists / Institute of Education Sciences
Bridget	Cherry	Founder and CEO	Cherry Strategies

Data Science

First Name	Last Name	Title	Organization
Justin	Ballenger	Assistant Professor of STEM Education	Division of Professional and Continuing Studies, Morehouse College
Christopher	Brooks	Assistant Professor	School of Information, University of Michigan
Zarek	Drozda	Director	Data Science for Everyone
Kathi	Fisler	Founder	Bootstrap
Bill	Howe	Associate Professor, Information School	University of Washington
Shaun	Kellog	Interim Executive Director	Friday Institute for Educational Innovation
Nancy	Lue	Senior Director	Valhalla Foundation
Suyen	Machado	Director, Introduction To Data Science Project	UCLA
Karina	Montilla Edmonds	SVP, Head of Academies and University Alliances	SAP
George	Siemens	Professor and the Executive Director	Learning Innovation and Networked Knowledge Research Lab, University of Texas, Arlington
Jim	Stigler	Distinguished Professor of Psychology	University of California Los Angeles
Chéla	Wallace	Director of Science, Technology and Engineering	KIPP NYC

Appendices

Appendix A: Participant Profiles

Transforming Educational Technology Through Convergence
Participant List

Math

First Name	Last Name	Title	Organization
Ivón	Arroyo	Associate Professor, Manning College of Information and Computer Sciences	University of Massachusetts Amherst
Gabe	Cartagena	Director of Secondary Math	DC Public Schools
Kristen	DiCerbo	Chief Learning Officer	Khan Academy
Neil	Heffernan	William Smith Dean's Professor of Computer Science	Worcester Polytechnic Institute
Lewis	Leiboh	Senior Program Officer, EdTech and K12	Bill & Melinda Gates Foundation
Alexandra	Martinez	Instructional Coordinator, PK-12 Mathematics	San Diego USD
Erin	Ottmar	Associate Professor	Worcester Polytechnic Institute
Catherin	Paolucci	Assistant Professor, College of Education	University of Florida
Alfons	Prince	Middle School Mathematics Teacher	Center City Public Charter Schools
Steve	Ritter	Founder and Chief Scientist	Carnegie Learning
Sheela	Sethuraman	Founder & CEO	CueThink
Walter	Leite	Professor of Research and Evaluation Methodology	University of Florida
Kripa	Sundar	Founder and Lead Consultant	Edtech Recharge

Supporting Members

First Name	Last Name	Title	Organization
Ryan	Baker	Professor, Center Director	University of Pennsylvania, Penn Center for Learning Analytics
Ulrich	Boser	Founder and CEO	The Learning Agency
Lizzie	Jones	Program Manager	The Learning Agency
Seiyon	Lee	Graduate Assistant	University of Pennsylvania
Jaclyn	Ocumpaugh	Associate Director	University of Pennsylvania, Penn Center for Learning Analytics
Malia	Sieve	Consultant	The Learning Agency
Yeeva	Cheng	Program Associate	The Learning Agency

Appendix B: Meeting 2 Guide

Transforming Educational Technology Through Convergence
Meeting 2 - Subgroup Ideating October 2022

Meeting Guide
Greetings! And thank you, again, for lending your time, expertise, and creativity to ideating on how NSF could invest in transforming educational technology through convergence.

Between now and October 26, please hold your small-group meeting, write up answers to the question below, and send them back to Lizzie Jones at lizzie@the-learning-agency.com. We will then send you the ideas generated by other small-groups on October 28 for your feedback. We will use these answers and comments to structure the whole-group meeting on November 4, so please have your comments ready by November 1!

In your small-group meeting, please answer this overarching question:

What breakthrough education technology deliverable(s) could be developed over a two-year period that would excite the field and have the potential to make a large-scale, societal impact in data science education?

Remember, our current goal is to demonstrate to NSF that **Transforming Educational Technologies Through Convergence** is a viable track that they should select in the next phase of their Convergence Accelerator program.

In order to justify selecting this track, the NSF would like to see the potential for several deliverables in the "minimum viable product stage" within 2-3 years, that could result in large-scale, societal impacts within 5-10 years. Therefore, you and your subgroup's task is to ideate deliverables within your track.

Education has never been identified as a track, and now is our time!

Instructions
- Consider the questions below. We've provided space for five deliverables, but we're more than happy to receive more!
- Use whatever format is easiest for you (e.g., bullet form, full sentences). What's important is that we are able to understand what you're conveying, not the format in which it's provided.
- Please meet at least once synchronously and provide these responses to us by October 26. Then be ready to comment on ideas from other groups!

Questions to Answer
(Respond to the following questions PER DELIVERABLE. Once you have responded to these questions for each of your deliverable suggestions, please return/email this document to Lizzie Jones at lizzie@the-learning-agency.com.)

Subgroup Members: (list members below)
●
●
●
●

DELIVERABLE 1:

1. What problem are you seeking to address? (within your track's overall mission)

2. What deliverable are you proposing?

3. What specific outcome improvements could we expect with this deliverable?

4. How will this deliverable help promote equity?

5. What steps would developers need to take to ensure that this new deliverable is not introducing or reinforcing bias?

6. What disciplines and types of expertise are needed for the development of the solution?

7. What evidence can we give NSF that this deliverable could be created in 2-3 years? (be brief)

8. If a “minimum viable product” was ready in 2025, what societal impacts could we expect by 2030? Or by 2035? How could this transform the field for education for students, teachers, and/or researchers? Is there any evidence that can give NSF evidence of this? (be brief)

9. What else do you want to add about this deliverable?

DELIVERABLE 2:

1. What problem are you seeking to address? (within your track’s overall mission)

2. What deliverable are you proposing?

3. What specific outcome improvements could we expect with this deliverable?

4. How will this deliverable help promote equity?

5. What steps would developers need to take to ensure that this new deliverable is not introducing or reinforcing bias?

6. What disciplines and types of expertise are needed for the development of the solution?

7. What evidence can we give NSF that this deliverable could be created in 2-3 years? (be brief)

8. If a “minimum viable product” was ready in 2025, what societal impacts could we expect by 2030? Or by 2035? How could this transform the field for education for students, teachers, and/or researchers? Is there any evidence that can give NSF evidence of this? (be brief)

9. What else do you want to add about this deliverable?

DELIVERABLE 3:

1. What problem are you seeking to address? (within your track’s overall mission)

2. What deliverable are you proposing?

3. What specific outcome improvements could we expect with this deliverable?

4. How will this deliverable help promote equity?

5. What steps would developers need to take to ensure that this new deliverable is not introducing or reinforcing bias?

6. What disciplines and types of expertise are needed for the development of the solution?

7. What evidence can we give NSF that this deliverable could be created in 2-3 years? (be brief)

8. If a “minimum viable product” was ready in 2025, what societal impacts could we expect by 2030? Or by 2035? How could this transform the field for education for students, teachers, and/or researchers? Is there any evidence that can give NSF evidence of this? (be brief)

9. What else do you want to add about this deliverable?

DELIVERABLE 4:

1. What problem are you seeking to address? (within your track’s overall mission)

2. What deliverable are you proposing?

3. What specific outcome improvements could we expect with this deliverable?

4. How will this deliverable help promote equity? ç

5. What steps would developers need to take to ensure that this new deliverable is not introducing or reinforcing bias? 6. What disciplines and types of expertise are needed for the development of the solution?

7. What evidence can we give NSF that this deliverable could be created in 2-3 years? (be brief)

8. If a “minimum viable product” was ready in 2025, what societal impacts could we expect by 2030? Or by 2035? How could this transform the field for education for students, teachers, and/or researchers? Is there any evidence that can give NSF evidence of this? (be brief)

9. What else do you want to add about this deliverable?

DELIVERABLE 5:

1. What problem are you seeking to address? (within your track’s overall mission)

2. What deliverable are you proposing?

3. What specific outcome improvements could we expect with this deliverable?

4. How will this deliverable help promote equity?

5. What steps would developers need to take to ensure that this new deliverable is not introducing or reinforcing bias?

6. What disciplines and types of expertise are needed for the development of the solution?

7. What evidence can we give NSF that this deliverable could be created in 2-3 years? (be brief)

8. If a “minimum viable product” was ready in 2025, what societal impacts could we expect by 2030? Or by 2035? How could this transform the field for education for students, teachers, and/or researchers? Is there any evidence that can give NSF evidence of this? (be brief)

9. What else do you want to add about this deliverable?

Considerations During Idea Generation
(These are just questions to help guide the generation of ideas; you do not need to provide written answers to these questions)

Does this approach:
- Produce more actionable information about students and help develop meaningfully personalized instruction?
- Contribute to improving interventions to support students?
- Lower the burden on overworked teachers?
- Improve outcomes for underfunded schools?
- Increase teacher knowledge of the domain, technology, or pedagogical strategies?
- Help the field promote equity? Does it reduce biases that impact learners, whether coming from technology or human decisions?
- Help diffuse insights more widely across digital learning platforms, lowering the development burdens by reducing the cost of adapting processes to new contexts?

Appendix C: Feedback exercise document

Deliverable Feedback Protocol

The consultancy exercise is structured to help teams think expansively about a particular, concrete dilemma. A dilemma is a puzzle: an issue that raises questions, an idea that seems to have conceptual gaps, or something about a process or strategy that you just can't figure out.

This exercise should take approximately 30 minutes per presentation/discussion. Once the full exercise described below is complete, then switch roles with the other team. In the time allotted for this segment, you should be able to complete two discussions.

Getting started:

- Decide which team will be the first Presenting Team; the other will be the first Advisory Team.
- The Presenting Team selects a presenter and a notetaker. The notetaker captures the questions and responses throughout the exercise.
- The Advisory Team selects a timer to track time throughout the exercise.
- Begin the exercise!

Exercise:

1. Initial Presentation of the Challenge [3 minutes]: Presenting Team One member of the Presenting Team presents a quick overview of the sub-problem the team is trying to address within the track, and frames focused questions for the Advisory Team to consider.

2. Clarifying Questions [5 minutes]: Advisory Team Advisory Team members ask questions of the presenters that have factual answers of a phrase or two in length. They ask the presenters "who, what, where, when and how much" questions. Clarifying questions do not include "why?" or "what other approaches have you considered?" questions. The purpose of clarifying questions is to help the questioner better understand the presenters' situation. The notetaker writes down all these questions.

3. Probing Questions [5 minutes]: Advisory Team Advisory group members ask questions of the presenters that help the presenters clarify and extend their own thinking about the matter they have presented to the group. The group asks open-ended questions such as: "why...?" "what other approaches have you considered regarding...?"or "what do you think would happen if...?". The notetaker writes down all these questions.

4. Advisory Team Discussion [7 minutes] The Advisory Team members talk with each other while the presenters listen and take notes; the presenters are not allowed to speak at this time (except to answer a clarifying question if one arises). The Presenting Team turns off their cameras and attends to listening and notetaking without providing any kind of response to the speakers. This separation often feels awkward but it is only for a few minutes and the benefits can be substantial. Advisory team members aim to discuss the situation and possible ideas about solutions. It is important for the presenters to listen carefully and in a non-defensive manner. The notetaker writes down all these notes.

5. Presenting Team Response into Open Discussion [5 minutes] The Presenting Team responds to what the Advisory Team said in the previous section. The purpose of this section is not for the presenters to respond to everything the response group members said. Instead, the purpose is for the presenters to talk about what they heard that was most important to them, and any thoughts or questions that were stimulated by the group discussion. Once the presenters have responded to their own satisfaction and wish to engage in a more free-flowing dialogue, they indicate so to the group by explicitly saying that they are ready to discuss additional comments, ideas and questions. The notetaker writes down all these responses.

6. Adjustments [5 minutes] The Presenting Team discusses and the notetaker records any adjustments to the recommendation.

Once complete, the Presenting Team and the Advisory Team switch roles and repeat the exercise.