Extended Abstract for the
Symposium on Integrating AI and Data Science into School Education Across Disciplines 2025

COMPETENCIES AND CURRICULUM IMPLICATIONS AT THE INTERSECTION OF MATHEMATICS, DATA SCIENCE AND STATISTICS

Cathy Smith[1] & Vinay Kathotia[2]
[1]The Open University, United Kingdom, cathy.smith@open.ac.uk
[2]The Open University, United Kingdom

Focus Topics: AI and data science curricula and implementation in schools, AI and data science competencies

Context

The school curriculum in England is currently under broad review, with the 2024 change of government reopening debates about how the mathematical needs of all students are addressed. The UK Royal Society has contributed significant thinking in this area through its 4-year Mathematical Futures Programme, which proposed a new vision of mathematical education that anticipates and supports the role of mathematics for individuals, economies and society, strengthens diversity and reduces inequity (Royal Society, 2023). The work we report here was commissioned as part of this programme (Smith et al., 2023) in order to investigate possibilities of bringing data science into a mathematics curriculum and to suggest competencies for a range of school leavers.

English students currently study mathematics to 16, after which 30% of the cohort retake that course, 16.5% continue with university-track mathematics courses and the rest study no named mathematics. Statistics is an established part of the 5-16 mathematics curriculum, and English students achieve well on uncertainty and data test questions (Ingram et al., 2023) but there has been concern among statisticians that what is learnt in school – characterised by no use of technology, short closed questions and small data sets – is increasingly unlike statistics used at university or in employment (Davies & Sheldon, 2021). The growing international interest in data science education offers an opportunity to critique our curriculum.

Approach

In stage 1, the project team undertook a non-systematic review of English-language research literature comparing visions for mathematics, data science, statistics, and computing education (MDSC). Through snowball sampling we talked to international experts about mathematics curricula which already integrate aspects of data science, such as those of Aotearoa New Zealand and Ontario, and explored curricular initiatives including Introduction to Data Science (Gould, 2021) and Germany's ProDaBi: Project Data Science and Big Data in School (Biehler & Fleischer, 2021), which brings in elements of machine learning. This stage was guided by three questions:

What does the existing vision literature in MDSC indicate are areas of consensus *within* and *between* the relevant disciplines?

What are examples of stable, well-documented and high-performing curricula or curriculum initiatives considered to address these intersections of MDSC?

What are the implications for competencies for citizens and school curriculum models?

In stage 2, we drafted a framework of school leaver competencies at the intersection of mathematics, statistics, data science and computing and explored some systemic implications for England. These have since been developed by others into data skills for end of primary and lower secondary school (Jaques & Joubert, 2024). More generally, our review led us to consider how a curricular move towards data science, characterised by statistical and quantitative reasoning plus a central role for technology, has the potential to change elements of how secondary mathematics is taught and learnt, and challenge aspects of its presentation as a discipline.

Discussion

Data scientists and statisticians will not be surprised that a strong shared emphasis in the literature we read was the overarching activity of posing and answering questions in context. The recent visions for computing education also emphasise that computational methods are designed in context, for a purpose and for people, with Tedre and Denning (2016) tracing this shift to the gradual alignment

In S. Podworny & S. Schönbrodt (Eds.), *Towards Fostering AI and Data Science Literacy in Schools Across Disciplines. Proceedings of the 1st Symposium on Integrating AI and Data Science into School Education Across Disciplines (AIDEA1 2025)*, Salzburg, Austria.

of computing with engineering rather than mathematics. This area of consensus does agree with international curricular approaches to mathematics, such as the PISA's mathematical reasoning cycle Formulate – Employ - Interpret and Evaluate, but it stands in contrast with the 2017 English mathematics curriculum which makes no mention of questions having source or purpose. Previous versions have included a data handling cycle for statistics, and thus we considered it reasonable to propose that the intersections of MDSC could be structured around a central competence of conducting iterative and holistic enquiry concerned with the study of variation and generalisability, with precision and using technology, attending to context and ethics.

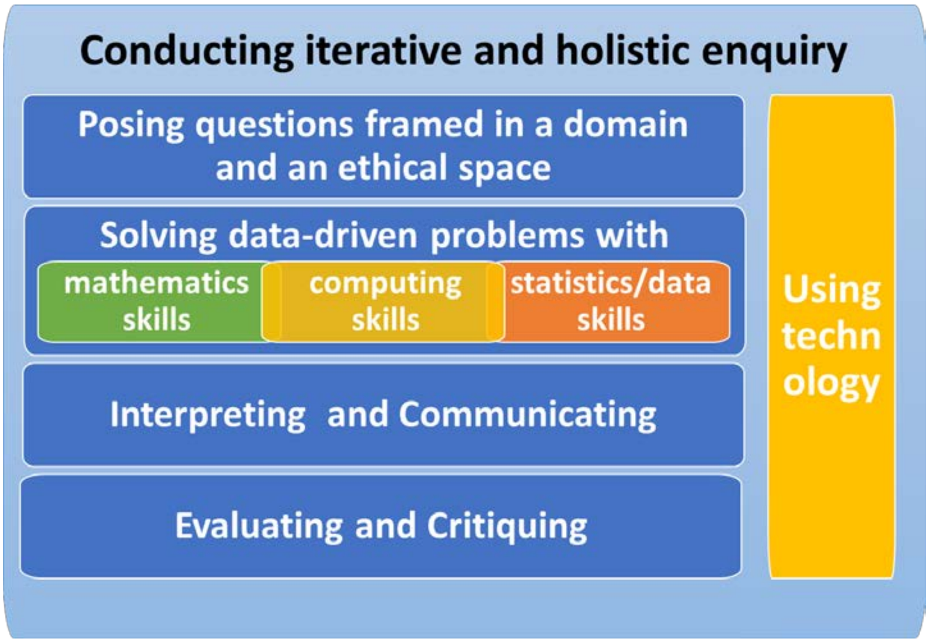

Figure 1: Mathematical, computational and data literacies as posing and answering questions

The six high-level competencies shown in Figure 1 are common across the literature reviewed. They are related but different, and each is necessary for the resulting activity to be considered as characteristic of the intersection of mathematics, statistics data science and computing (Smith et al., 2023, p. 25).

Where our conversations with data scientists challenged the familiar use of that statistical cycle is that the availability of multi-variable data sets may well precede, or even inspire, questions about the relevant domain. There are further elements of this central competence that change elements of how secondary mathematics is taught and learnt.

Context

Context figures throughout statistical inquiry: it speaks to the provenance of data, to the appropriateness of different questions, as well as to the value and applicability of any conclusions. Similarly, the tools – digital or otherwise - relevant to each context are so important to statistics/data science that they cannot be omitted from the intersection. This is notably different to common forms of mathematical thinking which instead valorise abstraction and thus, arguably, an early move away from context which is treated as obscuring structure (Cobb & Moore, 1997) and from artefact-based to symbolic representations. This necessitates a different and plural aesthetic of mathematics.

Ethics

Mathematics has had a complicated relationship with ethics, positioned by many teachers as an abstract field aside from ethical concern yet, functioning as a technology with moral and political effects, it is not neutral. The character and practice of data science then arguably have the potential, not to introduce ethics into mathematics, but to make ethical components more visible and more personal to pupils and teachers, giving learners opportunities to raise and answer ethical questions *with* data science and AI tools, and also *about* the use of such tools.

Mathematics for data science

There is some consensus about areas of mathematical activity in the secondary curriculum relevant for solving data-driven problems. From our reading, we identify: using quantities and methods, analysing covariation, reasoning mathematically, using and comparing representations (e.g. 2-way tables, schematics, graphs, symbolic forms), and using aids and tools. We note also that spatial reasoning and visualisation indirectly underpins graphical and data displays.

We developed the overarching and MDSC-specific competencies in three layers of increasing specialisation. Firstly, the core competencies **for all** school leavers, required for active citizenship and employment, who need to understand and respond critically to numerical information and claims and to use these for decision-making. The next layer of competencies are **for many** students (future practitioners), who may need to use data skills flexibly in a range of contexts and/or have additional responsibilities for choosing and critiquing data, models and tools. Finally, **for specialists**, who would be extending disciplinary knowledge and developing new analytic models.

Implications for curriculum design and implementation

We briefly look at how shifts in practice entailed by a data-centred framing could play out in data science and mathematical competencies and implications for curricular design and implementation.

Strong data stewardship is integral to data science, from attending to data privacy, security and fair usage for all, to specialists being able to enact data handling practices that support reproducibility, reusability and ethical use of data. Attending to ethics in data science moves beyond the good data practice in current school curricula (e.g. looking out for misleading graphs or not mistaking correlation for causation). It considers more fundamental issues of biases and fairness in the choice of data, tools and representations. Such practice, e.g. examining algorithms for their opacity and potential for harm, can extend beyond data science to mathematics, with learners being supported in exploring the wider impact of, say, maximising for profit or considering the sustainability of exponential growth models.

This questioning of mathematics, alongside using mathematics for questioning, can be seen as part of a significant shift to a more holistic, critical and nuanced use of mathematics, where, to quote Tukey (1962), "Data analysis must use mathematical argument and mathematical results as bases for judgment rather than as bases for proof or stamps of validity." (ibid, p. 6). The resulting greater role for uncertainty and use of context within mathematics, and the centrality of questioning, or even critical questioning (Kathotia et al., 2025), is not without its challenges, especially in shifting norms, expectations and sociocultural practices in mathematics classrooms. This may require related teacher professional development, in particular, support for mathematics teachers to explore more open-ended discussion-based approaches as in the social sciences. Data-centred problem solving could provide an opportunity for cross-disciplinary collaborative work, where learners and teachers explore the contrasts and parallels between different disciplinary approaches to questioning, classroom discussion, and notions of rigour.

Reflections following AIDEA 2025

The AIDEA symposium drew our attention to the many and differing types of association between Mathematics, Data Science, Statistics, Computing (MDSC) and Artificial Intelligence (AI), and the affordances and challenges of attending to these finer-grained intersections and differences.

AI

Our report for the Royal Society (Smith et al., 2023) did not go into AI or Generative AI (though there was some reference to machine learning, e.g. via ProDaBi (Biehler & Fleischer, 2021)) and the symposium provided an opportunity to consider related overlaps and differences between MDSC and AI education. One key point that emerged for us was that a framing of critical questioning that we are proposing to bring together MDSC at school level (Kathotia et al., 2025) would still be meaningful for engaging with AI and Generative AI in schools.

Separating Statistics and Data Skills

While we had pooled statistics and data science skills (Figure 1) given their substantial overlap, naming and drawing attention to the processes of working with data, e.g. data cleaning, data moves (Erickson et al., 2019), data workflow, data visualisation, could help draw attention to specific skills and the key role of data in the utility, validity and trustability of data scientific and statistical approaches.

Statistical literacy

Equally, Iddo Gal in the final AIDEA plenary session (Gal, 2025) drew attention to how few of us use or manipulate actual data sets in our lives and work, yet we are aware of and use statistical measures and indices, especially when consuming news and services. Working with such measures to foster "critical/smart consumers" could help develop "statistical, data science and AI literacy" (Gal, 2025) **for all**, as proposed in our competencies. We agree that school data science does require statistical literacy and should coordinate with related mathematical underpinnings, for example in understanding the implications of different ways of constructing such measures. Further, much of being able to evaluate information and findings involves working with text, and we should not lose sight of supporting basic literacy if we want to foster data and mathematical literacy.

Reading and Writing

The tensions between our roles as producers and consumers of MDSC was highlighted by some researchers' use of Bloom's taxonomy, or variations (Krathwohl, 2002), to place creation at the apex of a hierarchy of skills for learners, whereas others saw opportunities for learners at all levels to engage with reading, applying, analysing, evaluating and creating in the context of MDSC. These roles also have contextual and ethical implications given opportunities for 'reading and writing the word and the world' (Weiland, 2025), as well as pedagogical affordances given opportunities for learners to see themselves as creators and see themselves and their interests in the data (Tedre et al., 2025).

Mathematics for AI and AI for mathematics

A number of presenters discussed the mathematics needed for machine learning and AI (Bata et al. (2025), Frank (2025)), and why some of this should be in the school curriculum, not just for future creators of technology but to demystify AI for all students. AI constructs and concepts can also serve as contexts for fundamental mathematical concepts, e.g. distance, circles, area, density, dimensions, tree diagrams, two-way tables, optimisation, etc. Thus, while there are many unresolved questions as to how we should structure data education, AI serving as context for established parts of the mathematics curriculum provides an opportunity to start on MDSC convergence and coherence before we address more complex issues around the positioning and delivery of data education.

Curricular development and delivery

There is a fair amount of work afoot on developing competencies and learning progressions for data and AI education, including no-coding approaches. A common challenge (across continents) is around how and where in the curriculum we place it, how we support teachers to deliver interdisciplinary content that attends to context and ethics, and how we do this equitably.

References

Bata, K., & Frank, M. (2025, February 24–28). *Unsupervised machine learning as learning content in lower secondary school* [Symposium presentation abstract]. Symposium on Integrating AI and Data Science into School Education Across Disciplines, Salzburg, Austria.

Biehler, R., & Fleischer, Y. (2021). Introducing students to machine learning with decision trees using CODAP and Jupyter Notebooks. *Teaching Statistics, 43*(S1). https://doi.org/10.1111/test.12279

Cobb, G. W., & Moore, D. S. (1997). Mathematics, Statistics, and Teaching. *The American Mathematical Monthly, 104*(9), 801. https://doi.org/10.2307/2975286

Davies, N., & Sheldon, N. (2021). Teaching statistics and data science in England's schools. *Teaching Statistics, 43*(S1), S52–S70. https://doi.org/10.1111/test.12276

Erickson, T., Wilkerson, M., Finzer, W., & Reichsman, F. (2019). Data Moves. *Technology Innovations in Statistics Education*, 12(1). http://dx.doi.org/10.5070/T5121038001 Retrieved from https://escholarship.org/uc/item/0mg8m7g6

Frank, M. (2025, February 24–28). *What university-level mathematics needed to understand deep learning can be made accessible in high school?* [Symposium presentation abstract]. Symposium on Integrating AI and Data Science into School Education Across Disciplines, Salzburg, Austria.

Gal, I. (2025, February 24–28). *Infusing 'statistical, data science and AI literacy' for the general learner population: conceptual, instructional, and systemic challenges* [Symposium presentation abstract]. Symposium on Integrating AI and Data Science into School Education Across Disciplines, Salzburg, Austria.

Gould, R. (2021). Toward data-scientific thinking. *Teaching Statistics*, *43*(S1). https://doi.org/10.1111/test.12267

Ingram, J., Stiff, J., Cadwallader, S., Lee, G., & Kayton, H. (2023). *PISA 2022: National report for England*. Department for Education. https://www.ora.ox.ac.uk/objects/uuid:f90b59cc-1b5c-4c53-bb1b-26e936c4449a

Jaques, L., & Joubert, M. (2024, November 16). *A vision for meaningful data education across the four nations of the UK*. The Joint Mathematical Council of the United Kingdom. https://www.jmc.org.uk/2024/11/16/a-vision-for-meaningful-data-education-across-the-four-nations-of-the-uk/

Kathotia, V., Smith, C., & Ward-Penny, R. (2025). *Data Science, Mathematics and the Cohesive Potential of a Critical Questioning Paradigm* [Manuscript submitted for publication].

Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice*, 41(4), 212-218. https://doi.org/10.1207/s15430421tip4104_2

Royal Society. (2023). *A new approach to mathematics and data education*. https://royalsociety.org/-/media/policy/projects/maths-futures/new-approach-to-mathematics-and-data-education.pdf?la=en-GB&hash=8FFA5FD36A6E5606C834AE38DB33B37C

Smith, C., Kathotia, V., Ward-Penny, R., Howson, O., & Wermelinger, M. (2023). *Mathematical and Data Literacy: Competencies and curriculum implications at the intersection of mathematics, data science, statistics and computing*. The Open University. https://royalsociety.org/-/media/policy/projects/maths-futures/intersection-mathematics-data-statistics-computing.pdf

Tedre, M., & Denning, P. J. (2016). The long quest for computational thinking. *Proceedings of the 16th Koli Calling International Conference on Computing Education Research*, 120–129. https://doi.org/10.1145/2999541.2999542

Tedre, M., & Vartiainen, H. (2025, February 24–28). *Generation AI: pedagogical frameworks and tools for AI education* [Symposium presentation abstract]. Symposium on Integrating AI and Data Science into School Education Across Disciplines, Salzburg, Austria.

Tukey, J. W. (1962). The Future of Data Analysis. *The Annals of Mathematical Statistics*, *33*(1), 1-67.

Weiland, T. (2025, February 24–28). *Research to practice: designing learning experiences for teachers around reading the word and the world with data visualizations* [Symposium presentation abstract]. Symposium on Integrating AI and Data Science into School Education Across Disciplines, Salzburg, Austria.