Chang, M. et al. (Eds.) (2019). Proceedings of the 27th International Conference on Computers in Education. Taiwan: Asia-Pacific Society for Computers in Education

A SySTEMic approach to Data Literacy

Jon MASON*, Khalid KHAN & Gregory SMITH
College of Education, Charles Darwin University, Australia
*jon.mason@cdu.edu.au

Abstract: We outline an emerging research agenda that brings together three distinct topics relevant to the changing requirements of school education: STEM education, inquiry learning, and data literacy. When examining current curriculum materials in Australia we find that data in the digital environment is primarily seen as something to be examined in terms of its representations, interpretations, and visualisation – all activities that can be understood as downstream from data production. We argue that in the same way that inquiry learning involves a shift from the teacher to the student as the pivotal questioner, that a comprehensive and systemic approach to understanding data in the digital environment must involve appreciation for its production as well as consumption. Teaching STEM as an integrated and interdisciplinary way to decomposing real-world problems, not just the STEM subjects, requires a systemic approach running through and across the curriculum. It also now demands data literacy.

Keywords: STEM, inquiry, questioning, questions, data literacy, curriculum

1. Introduction

A character in a 1936 story *Swadeshi Rail* by a famous Urdu writer, Shaukat Thanvi, inquires *which station is this train heading to?* He receives a response that *it depends upon the majority of passengers*. A Facebook post recently asked a similar question, *who decides what is fake?* Big data may provide a compelling new resource that potentially can be used to solve many real-world problems, but it presents similar challenging questions if we scrutinize the underpinning instruments that produce it. It would be a big mistake if the sheer volumes involved in big data led us to think such scale has equivalence to a comprehensive census.

Combining big data with algorithms enables decision-making processes to be automated; however, most people only experience algorithms as 'black boxes' and must either accept the data produced as valid or not. In practice, despite the growing public outcry regarding this issue, such data is usually attached to a service. A common example is banking, and the sources of data used to produce our credit rating. Another well-known example is the data collected via consumer loyalty cards. But what of the data collected that we don't really know about?

In previous work two of us defined proposed a definition of data literacy that extends beyond the learning of data through computational and statistical representations that also includes making sense of diverse data types, to recognise the stories that data visualizations can convey, and to acknowledge the role of the data storyteller (Khan & Mason, 2016). In our combined research agenda, we focus on the progression of understanding the nature of data from first encounters in educational activities that begin in early childhood through to senior schooling years. In doing so, we propose and advocate for a systemic approach to data literacy that aligns with a curriculum where science, technology, engineering and mathematics (STEM) is embedded, though not necessarily explicitly.

In 2018, we surveyed STEM experts from around the world attending three separate international conferences focused on computers in education. Among the findings are themes of concern that were identified in the literature a few decades ago, leaving us to conclude that systemic change of embedding of STEM in the curriculum has not yet taken place. Additionally, responses to the survey reveal that data literacy is yet to be recognised as a foundation that underpins STEM education.

2. Change of Millennium

Leading up to 2000 there was widespread public discourse concerning the school curriculum and how to prepare the next generations for a technology-transformed workplace. In Australia, the focus shifted to development of skills in cross-disciplinary, critical and creative thinking, problem-solving and digital technologies. These objectives are now central to the Australian national STEM school education strategy. It recognises the role of STEM at all levels of schooling (Education Council, 2015). In parallel, a global agenda focused on *21st century skills* has evolved (Griffin, McGaw, & Care, 2012). In the Australian Curriculum, this is embedded as ‘general capabilities’. Both agendas align with a shift toward student-centred pedagogies enabling self-directed learning.

In terms of calibrating teaching and learning to meet these new challenges we question whether the core competencies of creativity, critical thinking, collaboration, and communication are beginning to look too general? What about problem-design and data literacy? And, while computational thinking, design thinking, and systems thinking have all become embedded into the Australian curriculum, these modes of thinking are not enough in an environment where false information is prevalent and data visualization tools can be easily manipulated. Fake news and cyberwarfare are becoming increasingly sophisticated and both require a response that is acutely analytical, discerning and agile – and these abilities require questioning techniques that probe deeper than search. We need to focus curriculum on such questioning techniques and shift legacy and traditional pedagogies that focus on ‘thinking in answers’ to a mindset that is nurtured through the inquiry of ‘thinking in questions.’

3. STEM and Data Literacy

Abundance of data now is a feature of our times. Little data collected with rulers, thermometers and stop watches has morphed into big data (Lohr, 2012). Science has moved to using big data to investigate ‘big problems’; gene expression, plate tectonics, algal blooms, disease, and the fundamental particles of the universe to name a few. Science has entered a ‘fourth paradigm’ of data-intensive investigation (Tansley & Tolle, 2009). A key implication for inquiry is that it is now closely coupled with handling data – its production, processing, and interpretation. Thus, Earl and Katz (2006) have described data literacy as a skill that uses evidence to systematically consider an issue from a range of perspectives to explain, support and challenge established thinking (p.45).

Because data is produced in a diversity of ways across all disciplines a systemic way of dealing with it across the curriculum is needed. Table 1 illustrates how the number ‘5’ can be inferred (as data) in various ways.

Table 1

Representations of the number five

five	5	V	𝍸	2 + 3		

In primary schools, students often see the numeral five as a counting number, something to add, subtract, multiply and divide, but the meaning comes from a specific activity or context. Often the inquiry takes the fair test approach: change one variable, keep all other variables the same and measure the change. The measure skill moves from bigger-smaller to tools of measurement to discern how much bigger or smaller, so units become important. The journey of the numeral five in primary schools is again bound in the inquiry as the amount of change. The plants in the sun grow five centimeters more than those in the dark. Build a spaghetti bridge, a well-known STEM activity, one block breaks the bridge, another bridge holds five blocks, the data suggest that the bridge that held five blocks is stronger. As for the cars on the ramp: one block verses five blocks have now a measured height, a measured distance travelled, and for older primary students a speed calculated. Again, the data is generated through measurement to answer the inquiry question and visualised usually in a graph.

Secondary school brings, what Metz (2015) calls the ‘Cambrian explosion of data’. However, the inquiry is often the same process: question or problem; data generated, analysed, interpreted and visualized; more precise measurements, more variables, more data. The numeral five is again a measure of difference, yet it could be the slope of the line in a graph or the number of trials in an inquiry.

The Australian Curriculum already touches upon on issues concerning data usage though 'data production' itself is not yet explicit. Given that data can be produced automatically and intentionally within digital environments, data literacy requires attention to the full scope of data production.

We are now pursuing a core research question: *in what ways can we articulate a systemic approach to embedding data literacy within a STEM curriculum from early childhood to senior schooling?*

Acknowledgements

We thank Charles Darwin University for a grant (#H18098) that enabled this research.

References

Earl, L. and Katz, S. (2006), *Leading Schools in a Data Rich World*, Thousand Oaks, California.

Education Council (2015). *National STEM school education strategy 2016-2026.* Australian Government.

Griffin, P., McGaw, B., & Care, E. (Eds.). (2012). *Assessment and Teaching of 21st Century Skills.* Springer.

Khan, K., & Mason, J. (2016). Data, the Story, the Storyteller, in Chen, W. et al. (Eds.) *Workshop Proceedings of the 24th International Conference on Computers*, India: Asia-Pacific Society for Computers in Education. (2016). pp. 142-144.

Lohr, S. (2012). *Big Data and Big Brother*. New York Times: New York.

Mason, J., Khan, K., &. Smith, S. (2016). *"Literate, Numerate, Discriminate – Realigning 21st Century Skills",* in Chen, W. et al. (Eds.) *Proceedings of the 24th International Conference on Computers in Education*. India: Asia-Pacific Society for Computers in Education. pp. 609-614.

Metz, S. (2015). Big Data. *The Science Teacher, 82*(5), 6.

Tansley, S., & Tolle, K. M. (2009). *The fourth paradigm: data-intensive scientific discovery* (Vol. 1). A.J. Hey (Ed.). Redmond, WA: Microsoft research.

WORK-IN-PROGRESS POSTERS