

Teaching in a World of Messy Data

CHAD DORSEY

We live in a world transformed by data. This fact may have never been more apparent than across the past year as we have witnessed a global pandemic, economic shock waves, critical societal reckonings with questions of race and equity, and a national election, each inextricably intertwined with questions of understanding data. The time has come to ensure that learners of all ages can understand and work with data effectively, and that these opportunities are provided equitably to all, especially to underrepresented learners.

But what does it mean to prepare learners in this way, and how do we do it? These are key questions indeed. Certainly, what we're doing now is not sufficient. Even in science class, data still sit on the margins far too often. If learners are to gain the facility—and fluency—with data that the future demands, they must have regular, meaningful experiences with data throughout their learning. In recalling that data are *numbers with context*, we must recognize that learners gain understanding by working across subjects through experiences with authentic data encountered in contexts relevant to their own lived experience.

To provide what learners need, we must think differently. Preparing learners for a world drenched in data means changing how we teach and learn with data overall. While data education is in many ways still an emerging field, some guideposts can help us think about how to ensure powerful data learning experiences.

Move beyond simply interpreting data, and toward exploring it

Learners need to be literate with data. That term, however, calls to mind the

ability to simply *read and interpret* data. But learners need to be able to do much more than simply find values in a table or interpret a visualization handed to them. They must be able to create their own visualizations, exploring rich data sets until data itself becomes a medium in which they work fluently, with grace and mastery. Achieving this requires new thinking about learning—and the right technology tools. Our Common Online Data Analysis Platform (CO-DAP) is one such example, designed to enable low-threshold, high-ceiling exploration that empowers learners to take data into their own hands.

Support learners in making original discoveries in data

Learners must be able to work with data in ways that allow them to make their own discoveries freely and openly. This means ensuring that they have the freedom to do so, through work with data sets with multiple variables and multiple options, and with data they may often have generated themselves. This can require new approaches to teaching and learning that position learners to see graphs and data as *tools* rather than *tasks*. But such approaches, especially when they lead to higher-level investigation and are paired with mathematical modeling of physical situations, are fully in line with the *Next Generation Science Standards*.

Embrace messy data

Learners today need to work with data in ways that more closely reflect what they will encounter in their lives. Data professionals engage in robust data cycles that include discovery, wrangling, profiling, modeling, and reporting. Most importantly, real data is inherently messy. Grappling with errors, outliers,

and missing values is essential, as is experience with accessing data through technology and data moves such as transforming, filtering, or joining data sets. Of course, time is precious and challenges always need to match learners' abilities. But presenting only pre-sanitized data robs learners of important opportunities such as determining which attributes should be included or how they should be represented. We must embrace data's natural messiness in order to prepare learners for the key skills they will need to succeed in the future.

Empower learners to find their answer

When we empower learners with the right tools and give them sufficient agency, they become problem solvers. Offering learners the chance to generate and explore interesting, multivariable data means we can't predict precisely what they will find. But we do know that there's a much better chance that what they find, and the stories they tell with it, will be interesting to *them*. Such empowerment is at the heart of motivated and engaged learning, makes for rich, problem- and project-based learning, and forms the basis for equitable and culturally responsive pedagogical experiences.

Shifting our perspective is never easy, but this new vision is vital as we move into an unpredictable, data-filled future. By rethinking the ways we approach data in science classrooms and across the curriculum, we will create learners who are truly prepared for the complexities of data in work and life. ■

Chad Dorsey (cdorsey@concord.org) is president and CEO of the Concord Consortium, Concord, MA.