# Exploring a US Framework of learning progressions for k-12 data science education

Katherine M. Miller1 & Michelle Hoda Wilkerson2

1Concord Consortium, USA, kmiller@concord.org

2University of California, Berkeley, USA

Focus Topics: AI and Data Science Competencies

## Introduction

While data science education (DSE) is growing quickly as a field within educational research, it is still nascent, especially in regard to K-12 teaching and learning. This creates a troubling dichotomy as uptake of data science education work among K-12 practitioners and developers accelerates ahead of research in the field (Drozda et al., 2022; National Academies of Sciences Engineering and Medicine [NASEM], 2023). Without coherent research guidance, such implementation occurs with little understanding about what learners need to know and be able to do with data, and when within the course of their learning these aspects are appropriate (Israel-Fishelson et al., 2023; NASEM, 2023). As such, there is a need for a framework that conceptualizes data science learning at the K-12 level which could serve as a guide to policy makers, practitioners and researchers alike. In an attempt to build such a framework, the Concord Consortium and Data Science 4 Everyone joined together, with seed funding from the National Science Foundation and the Valhalla Foundation to facilitate a series of workshops across the field with the goal of building consensus Learning Progressions (LP) for K-12 DSE.

## Background

A recent U. S. National Academies workshop, Foundations of Data Science for Students in Grades K-12 established K-12 DSE as a crucial and growing field (NASEM, 2023). While calls for K-12 data science education research draw from a rich history of statistics education research including learning trajectory development particular to that field (e.g., Franklin & Bargagliotti, 2020), the National Academies workshop highlighted the need for bringing siloed fields of research together to build a more comprehensive, consensus understanding of the scope of data science education. The workshop framed the need for interdisciplinary research to inform areas as far-ranging as identifying the role for computational tools and technologies in data science education, understanding the needs for teacher preparation for data science education, answering questions about what data science education looks like in different school subjects and educational settings, and identifying and encouraging just and ethical approaches to data science education (NASEM, 2023).

Across a variety of international gatherings similar to the Foundations workshop (e.g., International Association for Statistics Education roundtable, International Society for the Learning Sciences workshop) participants, panelists, and co-chairs point to the need for better understanding of research-based learning progressions, identifying the importance of overall coordination and clarity across data science education research (Miller & Yoon, 2022; NASEM, 2023). We adopt the language of LPs because they can inform, and hold sway across multiple contexts including, but not limited to, state and national standards, large-scale assessments, and classroom practice (NSF, 2011). While this work will be sensitive to the strengths and weaknesses of LPs as identified in the research literature, it will be equally sensitive to the needs of the multiple audiences which it will be designed to serve. Given the nascent state of the field of DSE, many different groups stand to benefit from the resulting framework, and the potential applications to which it may apply comprise an equally broad set.

## Context and Framing

The effort to design consensus LPs for K-12 DSE began with an NSF funded workshop in October 2023. We brought together educational researchers, professional learning coaches, and practitioners to explore possibilities for the structure and functionality of LPs for DSE, as well as an initial attempt to map the content of those progressions. This workshop resulted in an initial high-level outline of strands of learning for DSE. Over the next several months, a series of design focus groups were held to gather expert input from a breadth of groups that represent the full spectrum of data theory and practice, spanning higher-education, industry, K-12 practitioners, and K-12 students themselves. These focus groups identified the knowledge, skills, dispositions, and critical thinking tools students should gain by the time they graduate as relates to data literacy, data analysis techniques, and other data-related technology. Each focus group created a prioritized list of competencies for data science, shared these lists with the larger community, and provided feedback on the initial high-level strands. The work of the focus groups led to a revision of the strands and additional ideas for the content and competencies that belong within each strand.

Using the results of the focus groups were brought to a writing committee consisting of experts from across the field who engaged in a three-day writing sprint workshop to draft progressions of competencies within identified strands and sub-strands and in alignment with previous work. These competencies were written through a focus within each strand and then compared across strands and grade bands to determine overlap and identify holes where content was missing. A small writing team from Data Science 4 Everyone and the Concord Consortium then engaged in an intensive editing process to align language, grain size, and focus across all strands, sub-strands, grade bands, and concepts.

## Framework for DSE Learning Progressions

The current iteration of the high-level framework for learning for DSE in K-12 includes five strands representing Data Literacy & Responsibility, Creation & Curation, Data Analysis & Techniques, Interpreting Problems & Results, and Visualization & Communication, each of which contains 3 to 5 sub-strands. Figure 1 depicts the strands, sub-strands, and descriptions of content that might fall within each sub-strand across grade bands. In addition to the five strands, four lenses will guide the writing and development of the competencies which make up the content of the LPs. These lenses are social and cultural impact, questioning and critiquing, the role of technology, including tools and AI, and the data cycle.

A data science information table

Description automatically generated with medium confidence

Figure 1: Framework for Data Science Education Learning Progressions for K-12

Within each sub-strand are 3 to 6 “concepts” which represent core ideas and practices within DSE that students should learn and engage with across their K-12 schooling. Each concept stretches across five grade bands (K-2, 3-5, 6-8, 9-10, 11-12) with one or more competencies identified within each grade band that build on one another to progress towards mastery of the concept by completion of high school. The current iteration of the DSE LPs are content agnostic with a focus on data concepts that stretch across disciplinary boundaries.

## Conclusion and Next Steps

A draft of the LPs will be released to the public as “under construction” by February of 2025. Once they are released, there will be multiple additional opportunities for different parties and communities of interest, as well as the wider public, to provide insights and feedback on the development, production and interactive structure of the framework. The work will proceed as an iterative design cycle, engaging interest groups in a sustained way going forward. In this way, the proposed activities will generate a framework clear and specific enough to inform work across both research and development, yet flexible enough to evolve and incorporate the many new findings certain to arise from each.

## References

Drozda, Z., Johnstone, D., & Van Horne, B. (2022). Previewing the national landscape of K-12 data science implementation. In Workshop on Foundations of Data Science for Students in Grades K-12.

Franklin, C., & Bargagliotti, A. (2020). Introducing GAISE II: A guideline for precollege statistics and data science education. Harvard Data Science Review, 2(4), 1-9.

Israel-Fishelson, R., Moon, P. F., Tabak, R., & Weintrop, D. (2023). Preparing students to meet their data: an evaluation of K-12 data science tools. Behaviour & Information Technology, 1-20.

Miller, K., Yoon, S. (June 2023). Supporting High School Science Teachers in Developing Pedagogical Content Knowledge for Data Literacy. Paper presented at The Annual Conference of the International Society of the Learning Sciences (ISLS), Montreal, QC.

National Academies of Sciences, Engineering, and Medicine. (2023). Foundations of Data Science for Students in Grades K-12: Proceedings of a Workshop. Washington D.C.: The National Academies Press.

National Science Foundation. (2011, July 26). Learning Progressions Footprint Conference. Learning Progressions Footprint Conference Final Report. Learning Progressions Footprint Conference, Washington, DC.