American Journal of Botany
BOTANICAL SOCIETY OF AMERICA
leading scientists and educators since 1893
NEWS & VIEWS

COMMENTARY

The growing importance of data literacy in life science education

J. Phil Gibson[1,3] and Teresa Mourad[2]

Manuscript received 20 May 2018; revision accepted 28 August 2018.
[1] University of Oklahoma, Norman, OK, USA
[2] Director of Education and Diversity Programs, Ecological Society of America, Washington, D.C., USA
[3] Author for correspondence (e-mail: jpgibson@ou.edu)
Citation: Gibson, J. P. and T. Mourad. 2018. The growing importance of data literacy in life science education. *American Journal of Botany* 105(12): 1–4.
doi:10.1002/ajb2.1195

Science is a data-driven process, and biologists use many types of data, ranging from genomic to geospatial in their research. Advances in research technology have rapidly increased the amount, diversity, and complexity of data available to researchers (Hampton et al., 2013). Likewise, tools to collect, analyze, and share data have never been more readily available due to similar progress in computer technology that have driven the "big data" era (Marx, 2013; Baker, 2017). With these unprecedented data collection and analysis capabilities, data skills are widely acknowledged to have an increasingly important role in the future workplace, and familiarity with data science will be a necessity (NAS, 2018). In addition to learning the core content and experimental methods of their field, science, technology, engineering, and mathematics (STEM) students at all levels also need to develop their data literacy (i.e., the ability to create and use different forms of data). More than just a skill to be learned, studies of agnotology have shown that data literacy is a key component of critical thinking that can protect against weaponization of data to sow doubt and cultivate misinformation about scientifically valid research while falsely elevating pseudoscience (Cook et al., 2014). Thus, to prepare students for the data-rich future that undoubtedly lies ahead, it is imperative that STEM educators rise to meet this challenge and promote the development of strong data literacy in our students (Donovan, 2008; Marx, 2013).

Because data are a core component of science, students need to develop a full appreciation of the significance of data as scientific evidence and understand how to correctly analyze and interpret it. To explore the current state and nature of data literacy in life science education, the Botanical Society of America, the Ecological Society of America, and the Society for the Study of Evolution (collectively the Life Discovery Partners) convened the 4th Life Discovery Conference at the University of Oklahoma under the theme "Data: Discover, Investigate, Inform". The conference brought together biology educators from high schools, higher education, and STEM education organizations to share data-focused, educational modules and discuss the best approaches to improving data literacy of STEM students. Conference attendees specifically addressed this latter point in a networking session focused on identifying specific characteristics of data literacy that we should be striving for in our classes and the best approaches to achieve them (http://www.esa.org/ldc/pastconferences/2017-conference/networking/). The central ideas, suggestions, and conclusions from the discussion are summarized here to stimulate individual reflection and promote further conversations on data literacy in biology education among colleagues, departments, and programs.

WHAT IS BIOLOGICAL DATA LITERACY?

Although the fundamental skills of using arithmetic operations, mathematical functions, modeling, and data presentation form its foundation, data literacy is much more than numeracy and graphing. Data literacy is a broader concept that also includes understanding (1) which quantitative, analytical tools to use; (2) how to apply these tools in a biological context as an integrated component of an experiment; (3) how to interpret analyzed data in light of a specific question or hypothesis; and (4) how to effectively communicate results in different platforms. Discussants identified key features of data literacy and the characteristic quantitative skills and conceptual understanding associated with different levels of data literacy for those features (Table 1). The different levels of data literacy, ability, and understanding in the matrix can be broadly applied to students at any educational level and can be a starting point to assist educators in identifying specific aspects of data literacy they want to emphasize in their learning outcomes and assessment (Hoffman et al., 2016).

HOW TO DEVELOP DATA LITERACY

While identifying components and levels of data literacy is relatively straightforward, approaches to developing these skills are less clear. Participants suggested two general areas that faculty can emphasize to have the largest overall beneficial effect in courses and across curricula (Fig. 1). First, the most widely shared opinion was that students must understand why they are collecting certain data and how those data will provide evidence to address the question they are exploring. Students at all levels are often adept at the process of collecting data, and numerous peer-reviewed, data-rich classroom and laboratory experiences are available through portals such as the Life DiscoveryEd Digital Library (https://lifediscoveryed.org). While these resources promote development of quantitative and experimental skills, it is imperative that educators also help students understand why these specific data are being collected and recognize how those data can provide appropriate evidence to answer a specific question. These additional steps are often overlooked or not given the full attention they deserve, but they are absolutely necessary to promote students' understanding of the fundamental importance of considering how the data they collect will be analyzed as an integrated, essential part of exploratory and hypothesis-driven scientific inquiry (Elliott et al., 2016).

Second, educators should provide frequent opportunities to develop specific skills and use of data technology to better communicate about data in science. Increased availability of computers in classrooms and laboratories has allowed instructors to move beyond dedicating valuable instructional time to lengthy calculations and instead place greater emphasis on giving students meaningful practice with data analysis and interpretation. Instructors should similarly provide multiple opportunities for students to use technology to produce graphics that effectively communicate the information in the data. A sentiment expressed by many participants was that students at all levels struggle with constructing and interpreting figures—a severe data literacy weakness. The group unanimously agreed that educators should provide repeated opportunities for students to explore presenting data using appropriate software to not only develop technical skills for figure and graph construction, but also to potentially improve their ability to interpret the diverse ways data are presented in figures they encounter (Maltese et al., 2015). Experience summarizing their data should help foster understanding of their research and a deeper understanding of how to interpret data presented by other researchers.

IMPLEMENTING STEPS TO DEVELOP DATA LITERACY

We live in a rapidly expanding universe of data collection and analysis possibilities. The task we face as educators is to determine effective ways to implement educational changes that will improve our students' data literacy and ability to succeed in this dynamic environment. Unfortunately, there is no simple, singular answer. Conference participants recommended that an initial step is for individual instructors to identify specific data skills or concepts to teach, decide how to balance teaching the methods of data collection with teaching methods of analysis, and integrate these skills and concepts into their courses and curricula (Speth et al., 2010). Having identified specific learning outcomes, instructors should then find or develop resources to achieve those outcomes and implement them in their courses. Although developing new resources can be a daunting task at times, educators should not feel compelled to create activities de novo. A wealth of peer-reviewed, data-focused modules representing a diversity of degrees of engagement is readily available and can be implemented almost immediately. Sites

TABLE 1. Matrix of different facets of data literacy and the features of understanding and competencies associated with different levels of data literacy for them.

Facets of data literacy	Basic data literacy	Intermediate data literacy	Advanced data literacy
Collection and recording	• Know how to use instrumentation and technology to collect and store data • Able to collect and record data accurately using technology	• Able to identify appropriate data to collect relative to a biological question and hypothesis • Know how to enter data into a spreadsheet or database	• Understand how to incorporate rigorous data collection and sampling methods into experimental design • Know how to store, manage, manipulate, or query a database
Calculation	• Know how to conduct mathematical calculations • Able to use spreadsheets and software to conduct calculations	• Understand the relationship between calculations and a biological question • Know how to apply mathematical tools and technology to conduct calculations with biological data	• Know how to choose appropriate mathematical tools for studies of biological systems • Understand how data are used to develop quantitative biological models
Analysis and interpretation	• Know how to describe data with statistics • Able to describe patterns in data	• Know how to analyze and interpret data using statistical tests • Be able to interpret results of statistical test relative to a biological question or hypothesis	• Understand how to incorporate data analysis and statistical methods into experimental design • Understand assumptions in analyses • Able to compare results among analyses
Communication	• Know how to use technology to construct tables and figures. • Able to describe graphical and tabular presentations of data	• Able to explain the relationships among data presented in figures and tables • Understand how to use data and analyses to argue from evidence	• Able to evaluate the strengths and limitations of their data and analyses • Understand the relationship between their data and other biological, scientific or societal issues

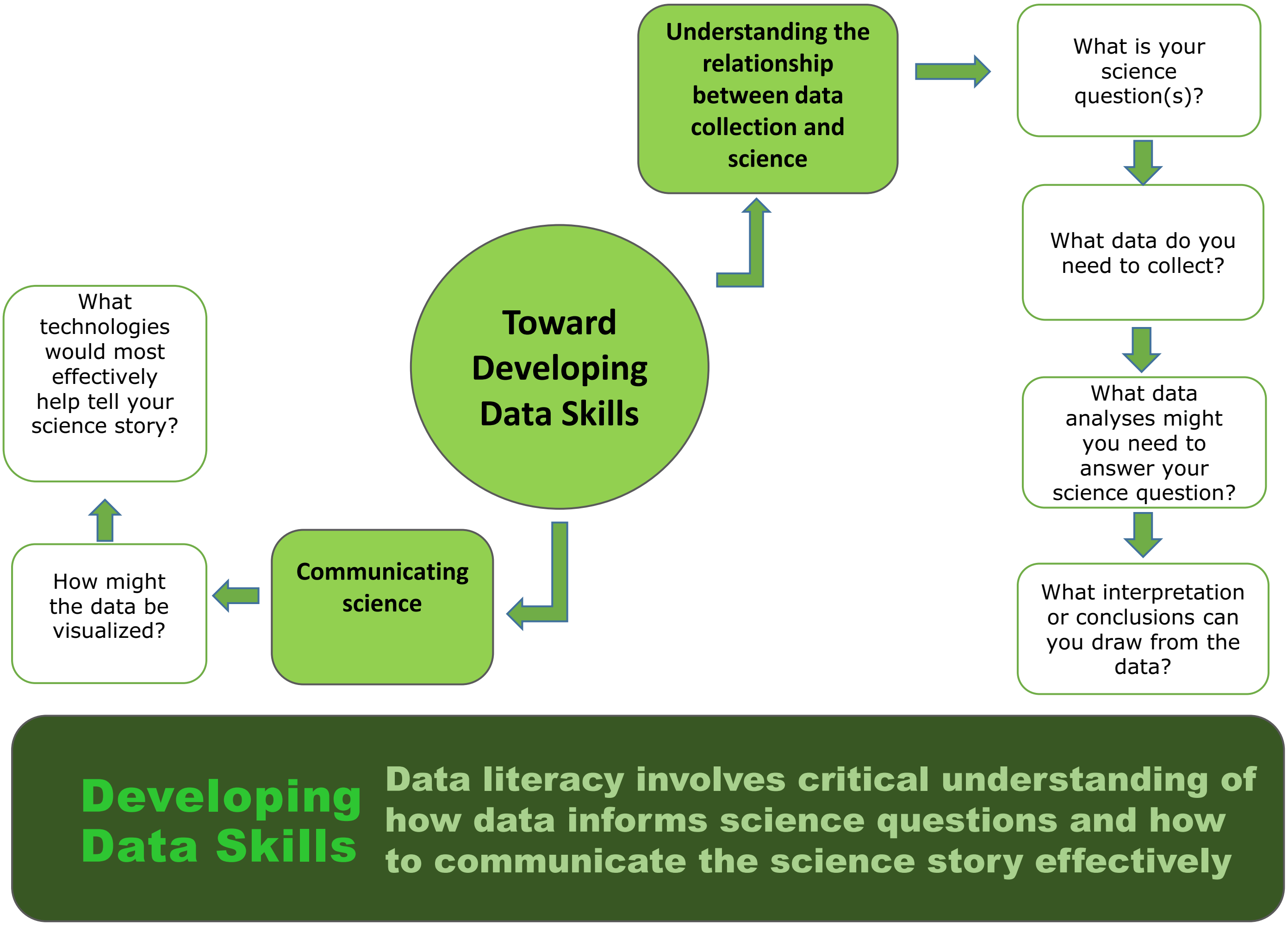

FIGURE 1. Suggested pathway toward developing the data collection and data communication skills that are essential to data literacy.

such as HHMI BioInteractive's Data Points (https://www.hhmi.org/biointeractive/data-points) or Data Nuggets (http://datanuggets.org) can be used to quickly bring data from the primary literature into the classroom to develop data analysis and interpretation skills. As mentioned previously, PlantED and other portals to the Life Discovery Education Library are sources of a wide variety of botanical, ecological, and evolution educational resources and the teaching guides to support their use. Teaching Issues and Experiments in Ecology, TIEE (http://tiee.esa.org/index.html), is another web-based collection of peer-reviewed teaching resources for STEM educators to explore. An even broader array of resources addressing different quantitative skills and resources for professional development of teachers are also available through the Quantitative Undergraduate Biology Education and Synthesis project (QUBES). The QUBES project not only provides access to peer-reviewed open educational resources (https://qubeshub.org/qubesresources), it also supports professional development for communities of educators who want to further develop their pedagogical skills through Faculty Mentoring Networks (FMNs). Participants who successfully complete FMNs sponsored by the BSA and ESA can earn designation as BSA and ESA Education Scholars (http://esa.org/fed/fmn/). After gaining familiarity with existing resources, discussion participants strongly encourage educators to develop their own unique resources and share them through these different venues so that the full tapestry of scientific inquiry is available for teachers to choose from and students to experience.

Achieving lasting change to students' data literacy cannot be achieved by instructors working alone, however, but will also require coordinated, sequential development of data literacy across the biology curriculum. It will be necessary for departments to identify, define, and articulate quantitative skills in their program's learning outcomes and coordinate the scaffolding of different aspects of data literacy and skills that will be implemented throughout their curriculum (Manduca and Mogk, 2002). A more ambitious suggestion in the discussion is for life science departments to begin communicating across STEM programs to further unify efforts to improve data literacy. We can achieve greater levels of success if we not only identify the different aspects of data literacy and skills that will be implemented throughout individual departmental curricula, but also communicate with other programs about how they are promoting data literacy in their fields (Manduca and Mogk, 2002). Because there is likely a great deal of overlap in desired learning outcomes, partnering in efforts to improve data literacy across STEM programs can only help further unify efforts to stimulate institutional change that improves critical quantitative abilities and data literacy in all students. Given that our goal as scientists is to use data to provide evidence-based explanations of the world we live in and provide answers to serious biological questions, efforts to improve data literacy in our students must be our highest priority. We encourage all of our colleagues to engage in discussions at all levels to consider how we can most effectively educate all of our students, STEM majors and nonmajors alike, to have the data literacy they will surely need to meet the challenges and changing face of science in the future.

ACKNOWLEDGEMENTS

We thank C. Adams, G. Hamerlinck, K. Jenkins, D. Lamar, J. Noor, P. Strode, and B. Wu for their assistance in facilitating the discussion at the LDC. We also thank P. Diggle and two anonymous reviewers for their comments on earlier versions of this manuscript. The LDC conference was supported in part by NSF Grant DUE #1742980 to the ESA (T. Mourad, PI; J. P. Gibson, co-PI).

LITERATURE CITED

Baker, B. 2017. Big data opens promising career paths for biologists. *BioScience* 67: 100.

Cook, J., D. Bedford, and S. Mandia. 2014. Raising climate literacy through addressing misinformation: Case studies in agnotology-based learning. *Journal of Geoscience Education* 62: 296–306.

Donovan, S. 2008. Big data: teaching must evolve to keep up with advances. *Nature* 455: 15260.

Elliott, K. C., K. S. Cheruveli, G. M. Montgomery, and P. A. Soranno. 2016. Conceptions of good science in our data-rich world. *BioScience* 66: 880–889.

Hampton, S. E., C. A. Strasser, J. J. Tewksbury, W. K. Gram, A. E. Budden, A. L. Batcheller, C. S. Duke, and J. H. Porter. 2013. Big data and the future of ecology. *Frontiers in Ecology and the Environment* 11: 156–162.

Hoffman, K., S. Leupen, K. Dowell, K. Kephart, and J. Leips, J. 2016. Development and assessment of modules to integrate quantitative skills in introductory biology courses. *CBE Life Sciences Education* 15: 1–12.

Maltese, A. V., J. A. Harsh, and D. Svetina. 2015. Data visualization literacy: investigating data interpretation along the novice–expert continuum. *Journal of College Science Teaching* 45: 84–90.

Manduca, C. A., and D. W. Mogk. 2002. Using data in undergraduate science classrooms. Science Education Resource Center, Northfield, MN, USA. Available at https://serc.carleton.edu/files/usingdata/UsingData.pdf.

Marx, V. 2013. Biology: the big challenges of big data. *Nature* 498: 255–260.

National Academies of Sciences, Engineering, and Medicine. 2018. Data science for undergraduates: opportunities and options. National Academies Press, Washington, D.C., USA. Available at https://doi.org/10.17226/25104.

Speth, E. B., J. L. Momsen, G. A. Moyerbrailean, D. Ebert-May, T. M. Long, S. Wyse, and D. Linton. 2010. 1, 2, 3, 4: infusing quantitative literacy into introductory biology. *CBE Life Sciences Education* 9: 323–332.