Data Literacy: Real-World Learning Through Problem-Solving With Data Sets

Author

Robin W. Erwin, Jr., Ph.D., is an Associate Professor and Chair for the Department of Professional Studies at Niagara University in Niagara, New York.

Abstract

The achievement of deep learning by secondary students requires teaching approaches that draw students into task commitment, integrated curricula, and analytical thinking. By using real-world data sets in project based instructional units, teachers can guide students in analyzing, interpreting, and reporting quantitative data. Working with authentic unfiltered, and previously un-interpreted data helps students to integrate data literacy with their disciplinary learning.

Keywords: data literacy, authentic learning, authentic data sets, performance based assessment, project learning, problem-based learning

There has been widespread agreement that students need to be highly proficient in literacy skills with language, but if students are to be well prepared for college and careers, they need to be "literate" in their abilities to work with data as well. In the face of increasing dependence on data analysis in all sectors of life, there is a growing call for students to develop data literacy (Gunter, 2007; Vahey et al., 2012). This article suggests that students can do so by participating in authentic data analysis tasks within the context of project-based learning (PBL) (Larmer & Mergendoller, 2010; Mergendoller, 2012; Moriyama, Satou, & King, 2002). In project based learning, students attempt to solve authentic problems in a process characterized by driving questions, inquiry and innovation, student voice and choice, and public presentation (Larmer & Mergendoller, 2010).

Students engaged in project based learning are expected to sustain their attention over large periods of time and to master content through an authentic experience that connects with multiple curriculum areas and relates to the learner's personal experiences. During that experience, teachers can cultivate higher order thinking by providing students with a data set in a raw form or by guiding students to create their own unique data set. Such data-analytic PBLs have several positive attributes. They are authentic learning tasks that encourage analytic thinking and meaningful experience with quantitative analysis, and they foster task commitment, interdisciplinary learning, and the integration of computers and literacy. Asking thick driving questions is a key to an effective PBL experience. To maximize the PBL experience, Mergendoller (2012) urged teachers to identify driving questions for a PBL experience that are not "Googleable," that is, not answerable by a simple Internet search.

Although it is important to work with real data sets, it is also true that the learning task and challenge must be developmentally appropriate (Wormeli, 2014), within what Vygotsky has named the zone of proximal development, so that the learner can perform the task if provided with peer or teacher support (Vygotsky, 1978). In that respect, a project based learning experience for middle grades would need to look and feel different from a PBL experience for late secondary grades, with learning goals and data set complexity that fit the developmental level of the learners. For example, a data set for middle grade students may have two or three arrays of simple data in columns, whereas a data set for upper high school may utilize more complex data in arrays of several columns, and may require more sophisticated types of statistical analysis.

Examples of Project Based Learning with Data Sets

In two separate six-lesson projects, I guided sixth to eighth grade parochial school students to use data sets obtained from www.gapminder.org. The first project was introduced as part of a study of safety issues and practice. The project examined mortality rates from injuries for children from birth to age 5 using a data set that included statistics from over 170 societies. The second project was part of a unit that was studying access to clean water resources around the world, especially in developing countries. The unit was already underway when a data-based project on the availability of improved water resources was introduced. The data set for that project spanned 10 years, and the analysis indicated improved water resources from more than 250 societies. During these data-based projects, I progressively instructed students how to:

1. Consider the "story" or guiding question for the data-centric unit
2. Learn about the power and utility of basic descriptive statistics for assigning meaning to large sets of data

3. Learn some basic skills for how to use spreadsheet software to examine large data sets
4. Examine a relevant data set
5. Clean the data in the data set
6. Analyze the data using basic statistical functions in a spreadsheet
7. Interpret the importance of the findings, and
8. Report the findings to an external audience.

Cleaning the data is a phrase to describe the work of improving the quality of data in a dataset. This is typically achieved by removing blank or duplicate rows, formatting numbers in a consistent manner, fixing discontinuities in column or row indexing, or any actions that improve the usability and quality of data in a dataset. As a result of these projects, the participating students demonstrated a satisfactory grasp of fundamental concepts of the scientific research process, the process or organizing and cleaning the data, the use of descriptive statistics to better understand the data, and the use of spreadsheet software to efficiently and accurately generate statistical analyses. The students also appreciated the greater objectivity gained through evidence when describing patterns in the data and reporting interpretations based on the data. There were rich conversations among groups of students that were challenged to analyze their data sets and develop inferences. The students, the classroom teacher, and the author considered the two projects to be valuable and successful learning experiences.

Research Supporting Data-based Projects

One scholarly initiative aimed at providing the resources and tools to support student learning with real data sets is the "Oceans of Data" project (Krumhansl et al., 2012). In this project, genuine scientific data sets and their graphic representations are key components in educating scientifically-minded learners. The project aims to maximize access to developmentally appropriate and useable data from oceanographic research across K-16 education. The graphic representations provided through this project are exemplars for visually communicating data analyses. Thorsen (2009) also described several examples of real data work using database and spreadsheets in which students were involved in collecting, managing, manipulating, interpreting, and applying data to real world problems.

Several researchers have documented some of the benefits of data based projects at the middle-school level. Vahey et al. (2012), conducted a study of seventh-grade students' work with data sets and essential questions. Students in this quasi-experimental study were taught to analyze authentic data sets by using proportional reasoning and then analyzing the data for what it could show regarding the driving questions. The students that were taught with the data-centric methods had significant learning advantages over the control group with a very large effect size for the treatment group. Hammond (2010) examined the positive interaction of learner cognition, motivation, and competence among middle grade students. The learning tasks in that study were highly authentic for the students, who served as apprentices in an exploratory, non-profit, educational, but real community bike shop.

At the college level, DeLuca and Lari (2011) conducted a study with students who were studying renewable energy sources by working with authentic, real world data to think and solve problems like real scientists. Findings in their study indicated that learners who accessed real-world data sets learned more thoroughly and with deeper mastery of scientific concepts than other learners who did not work with authentic data.

Wyner (2013) reported that teacher candidates who accessed and analyzed authentic climate data outperformed other groups of teacher candidates who pursued the same learning objectives in a book-centered curriculum. Ucar and Trundle (2011) observed the same benefits among teacher candidates who worked with authentic ocean tidal data retrieved from web databases; these learners gained a much stronger and more accurate grasp of the scientific principles involved in tidal dynamics than learners using other more traditional approaches.

In these examples, the students were functioning in the manner of neophyte scientists, collecting or acquiring authentic data relevant to their investigation, organizing and analyzing that data, drawing appropriate inferences, and reporting their finding to an interested audience. The data may be directly collected by the students or acquired as data sets available for scientific inquiry. The data might be as wide-ranging as historical documents, survey responses, field counts, or measures of performance.

Benefits of Projects Using Data Sets

The activities described above are clearly at the higher levels of Bloom's taxonomy (Anderson et al., 2001), and the highest levels of Webb's depth of knowledge criteria (Hess, Jones, Carlock, & Walkup, 2009). These and other benefits of projects using data sets are discussed in more detail below.

Authenticity Fosters Motivation and Task Commitment

Learning tasks are authentic when they have more than a grade value to the student, particularly when they have value in a community outside the classroom. The products of these authentic tasks are shared with more than just the teacher as evaluator, usually with some external audience. Authentic learning tasks focus on solutions to "real-world, complex problems by using role-playing exercises, problem-based activities, case studies, and participation in virtual communities of practice. Neo, Neo, and Tan (2012) argued for the importance of authenticity in learning as part of their study of student motivation in a web-based learning environment. They reported that when students perceived a learning task to be authentic, their commitment to the task increased. With careful planning and implementation, data-analytic PBLs may have this authentic quality for students.

For many students, real life data analysis and interpretation activities cultivate intrinsic motivation and sustained attention to the task. This sustained attention and intrinsic motivation is a significant benefit of real data since the learners' work toward the learning goals is typically difficult to achieve without such learner buy-in. Renzulli and Reis (1997) refer to this buy-in as "task commitment," and it is a key to effective learning in complex learning activities that require full concentration and effort over multiple work sessions.

Fostering Problem-Solving and Higher Level Thinking

The National Council of Teachers of Mathematics, in their *Principles and Standards for School Mathematics – Problem Solving* (2000), emphasizes the importance of the process of problem solving:

> Problem solving is an integral part of all mathematics learning. In everyday life and in the workplace, being able to solve problems can lead to great advantages. However, solving problems is not only a goal of learning mathematics but also a major means of doing so. Problem solving should not be an isolated part of the curriculum but should involve all Content Standards.

Problem solving means engaging in a task for which the solution is not known in advance. Good problem solvers have a "mathematical disposition", they analyze situations carefully in mathematical terms and naturally come to pose problems based on situations they see.

Working with students over an extended time to analyze real data, interpret, and apply the results fits well with the fourth and highest level of Webb's Depth of Knowledge criteria (Hess, Jones, Carlock, & Walkup, 2009), a level which is characterized by sustained effort over multiple learning sessions involving analytical thinking about complex problems.

Meeting Common Core State Standards

The Common Core State Standards argue for preparing students for college, careers, and other life activities. The Standards expect much in the way of analytical, multi-step thinking and problem-solving (Porter, McMaken, Hwang, & Yang, 2011). Working with raw data supports that kind of learning as learners clean and organize the data appropriately for analysis, apply the proper analytical tools, interpret the analysis, and make an appropriate application to a real world situation. People in data-related careers perform this organizing of raw data as part of their regular work. From this perspective, a teacher who starts with authentic data collection, or with data sets that are unrefined and un-interpreted, and then challenges learners to engage in this data-analytic work is helping students to reach the CCSS expectation of preparing for college and careers.

The Common Core State Standards (2010) include several Standards for Mathematical Practice that may be addressed in data-centric PBLs, including the following:

1. Make sense of problems and persevere in solving them.
2. Reason abstractly and quantitatively.
5. Use appropriate tools strategically.
7. Look for and make use of structure.

PBL using raw data may be designed to meet CCSS Anchor standards in other subjects such as Reading, and Speaking and Listening (see Table 1).

Table 1. ***CCSS Anchor Standards That Can be Addressed through PBL Using Raw Data***

CCSS Anchor Standards for Reading, grades 6-12

7. Integrate and evaluate content presented in diverse formats and media, including visually and quantitatively, as well as in words.
8. Delineate and evaluate the argument and specific claims in a text, including the validity of the reasoning as well as the relevance and sufficiency of the evidence.
10. Read and comprehend complex literary and informational texts independently and proficiently.

CCSS Anchor Standards for Writing, grades 6-12

7. Conduct short as well as more sustained research projects based on focused questions, demonstrating understanding of the subject under investigation.
8. Gather relevant information from multiple print and digital sources, assess the credibility and accuracy of each source, and integrate the information while avoiding plagiarism.
9. Draw evidence from literary or informational texts to support analysis, reflection, and research.

CCSS Anchor Standards for Speaking and Listening, grades 6–12

1. Prepare for and participate effectively in a range of conversations and collaborations with diverse partners, building on others' ideas and expressing their own clearly and persuasively.
2. Integrate and evaluate information presented in diverse media and formats, including visually, quantitatively, and orally.
4. Present information, findings, and supporting evidence such that listeners can follow the line of reasoning and the organization, development, and style are appropriate to task, purpose, and audience.
5. Make strategic use of digital media and visual displays of data to express information and enhance understanding of presentations.

Using Computer Software

Although data collection and analysis may be conducted without computer technology, students may benefit by using computer software as they work with real data. Programs like *Excel*, and in some cases *Systat* and *SPSS*, are within reach of most adolescent learners, especially when the students can see the power of this software to support efficient analysis of large data sets (Drier, Dawson, & Garofalo, 1999). Using such software also builds students' preparedness for college and careers. Too few educators appear to understand how important it is to engage students in these kinds of technology-supported higher order thinking activities (Lawrence, Calhoun, Milton, & Vozzo, 2013; So & Kim, 2009), but an increasing emphasis on data literacy in schools may be changing that.

For an excellent source of downloadable data sets on a variety of economic, political, and scientific issues for potential use in school settings, see http://www.gapminder.org/. Another useful source of authentic online data sets is https://tuvalabs.com/. *InspireData* is a software package developed by Inspiration Software, www.inspiration.com/InspireData, that supports students' data analysis, especially by creating visualizations of data for analysis. Another rich website for teachers seeking resources for data literacy is http://www.statlit.org/.

Conclusion

By using real data in Project-Based Learning, students will grow in numeracy and content knowledge, and they will be able to apply their literacy skill in disciplinary settings. Students will accomplish real intellectual work that is not just a practice activity within the school but rather has worth to the community. The integration of these multiple strands of the curriculum in data-centric PBLs provides students with meaningful study and learning in a potentially time-efficient manner.

Students engaged in analyzing data sets will broaden and deepen their knowledge and skills in content disciplines, thinking more like the experts in the discipline as they work in contexts that reinforce and reward expert and applied thinking. These students are more likely to experience satisfaction and even intellectual fulfillment when studying in the academic disciplines while pursuing these problem-solution tasks. Real data can contribute to real learning and true data literacy.

References

Anderson, L. W. (Ed.), Krathwohl, D. R. (Ed.), Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of Educational Objectives.* New York, NY: Longman.

Common Core State Standards Initiative (2010). *Mathematics standards*. Retrieved from http://www.corestandards.org/Math/.

Common Core State Standards Initiative. (2010). *The Common Core State Standards for English language arts & literacy in history/social studies, science, and technical subjects.* Retrieved January 5, 2015 from http://www.corestandards.org/ELA-Literacy/.

DeLuca, V. W. & Lari, N. (2011). The GRIDC project: Developing students' thinking skills in a data-rich environment. *Journal of Technology Education, 23*, 5-18.

Drier, H. S., Dawson, K., & Garofalo, J. (1999). Technology, mathematics, and interdisciplinary connections: Not your typical math class. *Educational Leadership, 56* (5), 21-25.

Gunter, G. A. (2007). Building student data literacy: An essential critical-thinking skill for the 21st century. *Multimedia & Internet@Schools, 14* (3), 24.

Hammond, C. F. (2010). *Envisioning competence: Learning, problem solving, and children at work in the exploratory bicycle shop* (Order No. 3444834). Available from ProQuest Dissertations & Theses Global. (859425675). Retrieved from http://ezproxy.niagara.edu/login?url=http://search.proquest.com/docview/859425675?accountid=28213.

Hess, K. K., Jones, B. S., Carlock, D., Walkup, J. R. (2009). *Cognitive rigor: Blending the strengths of Bloom's and Webb's Depth of Knowledge to enhance classroom-level processes.* (ERIC Document Reproduction Service No. ED517804)

Krumhansl, R., Peach, C., Foster, J., Busey, A., Baker, I., & DeLisi, J. (2012). *Visualizing oceans of data: Designing educational interfaces*. Waltham, MA: Education Development Center.Larmer, J., & Mergendoller, J. R. (2010). Seven essentials for project-based learning. Educational Leadership, 68(1), 34-37.

Lawrence, S. A., Calhoun, F., Milton, M., & Vozzo, L. (2013). Exploring teachers' perceptions of literacy and use of technology in classroom practice: Analysis of self-reported practice in one school district. *Journal of Literacy and Technology*, 14, 51-93.

Mergendoller, J. R. (2012). *Teaching critical thinking skills through project-based learning.* Retrieved from http://www.p21.org/news-events/p21blog/1097-teaching-critical-thinking-skills-through-project-based-learning.

Moriyama, J., Satou, M., & King, C. T. (2002). Problem-solving abilities produced in project based technology education. *Journal of Technology Studies, 28*, 154-158.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Retrieved from http://www.nctm.org/standards/content.aspx?id=26860

Neo, M., Neo, K. T-K., & Tan, H. Y-J. (2012). Applying authentic learning strategies in a multimedia and Web learning environment (MWLE): Malaysian students' perspective. *TOJET: The Turkish Online Journal of Educational Technology, 11* (3), 50-60.

Porter, A., McMaken, J., Hwang, J., & Yang, R. (2011). Common Core Standards: The new US intended curriculum. *Educational Researcher, 40* (3), 103-116.
Renzulli, J. S., & Reis, S. M. (1997). The school wide enrichment model: New directions for developing high-end learning. *Handbook of Gifted Education, 2*, 136-154.
So, Hyo-Jeong, & Kim, B. (2009). **Learning** about **Problem Based Learning**: Student teachers integrating technology, pedagogy and content knowledge. *Australasian Journal of Educational Technology, 25*, 101-116.
Thorsen, C. (2009). *TechTactics: Technology for teachers* (3rd ed.). Boston, MA: Pearson.
Ucar, S., & Trundle, K. C. (2011). Conducting guided inquiry in science classes using authentic, archived, Web-based data. *Computers & Education, 57*, 1571-1582.
Vahey, P., Rafanan, K., Patton, C., Swan, K., van't Hooft, M., Kratcoski, A., & Stanford, T. (2012). A cross-disciplinary approach to teaching data literacy and proportionality. *Educational Studies in Mathematics, 81*, 179–205.
Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.
Wormeli, R. (2014). Motivating young adolescents. *Educational Leadership*, *72* (1), 26-31.
Wyner, Y. (2013). A case study: Using authentic scientific data for teaching and learning of ecology. *Journal of College Science Teaching*, *42*(5), 54-60.