

IBI* SERIES WINNER

Engaging High School Students in Research on Smoking Behavior

Exploring Databases, an IBI prize-winning module, enables students to conduct epidemiological research on smoking behavior using a database.

Maureen M. Munn,^{1‡} Hiroki Oura,² Mark Gallivan,^{3†} Katie Van Horne,² Andrew W. Shouse²

Increasingly, scientists use information and communications technology to analyze large repositories of existing data. Engaging students in database investigations has great potential for providing authentic research experiences that are low cost and reflect contemporary science practice. Through a collaboration between Genome Sciences and the Institute for Science and Math Education at the University of Washington (UW), we developed Exploring Databases, a high school inquiry-based research project combining neurobiology, epidemiology, statistics, genetics, and database research to answer the question, “Why do some people smoke, and others don’t?”

Nicotine addiction remains the most common form of chemical dependence in the United States (1). Consequently, despite considerable public health investment, tobacco use is still the leading cause of preventable illness and death in the United States (2). Exploring Databases engages students in examining how environmental and genetic factors contribute to smoking addiction by using the Smoking Behavior database.

This database is the result of a previous science education project that involved high school students in planning and conducting a case control study that compared 300 adult smokers and nonsmokers (3). Research subjects completed a questionnaire regarding environmental influences on their smoking behavior. They also gave a small blood sample that was used to genotype their DNA at three candidate gene regions shown to be associated with smoking behavior: a dele-



Students and teachers engaged in their research. (Left) Students collaborate as they query the database. (Right) During workshops, teachers also work together to test their hypotheses.

tion in the promoter region of the dopamine receptor gene; a synonymous substitution in the dopamine receptor gene; and a substitution in an intron of the dopa decarboxylase gene (3). Questionnaire and genotyping data for each subject were entered into the database.

The Exploring Databases curriculum module consists of seven 1- to 2-hour lessons, including foundational activities and student-led investigations (see supplementary materials). The curriculum is taught by teachers who have attended a professional development workshop. In the first three lessons, students learn different aspects of human subjects research, discuss variation in smoking behavior, and study the biology of nicotine addiction. In lesson 4, students learn the case control study design and epidemiological analysis, including criteria used to distinguish causality from associations (4). Lesson 5 focuses on the fundamentals of statistics to estimate the strength and significance of associations. Throughout these lessons, students watch taped interviews of scientists in related fields and explore the role of databases in contemporary research. They develop an overarching hypothesis related to genetic or environmental influences on smoking behavior by reviewing profiles of smokers, examining published research and reflecting on their own experiences.

In lesson 6, student research teams identify questions that address their overarching hypothesis and use the database for hypothesis testing (see the first figure, left). The online database interface provides visual support to

guide students as they submit queries, estimate statistics, and interpret their results (see the second figure). Students then apply the criteria for causality to determine whether an association can be considered causal. In lesson 6, students are instructed to conduct, at most, four statistical tests related only to their overarching hypothesis to avoid false-positive results. However, in lesson 7, students “mine” the database by analyzing many questions and exposure combinations to generate a new hypothesis for a future hypothetical research study.

To conclude the module, students create a presentation in PowerPoint or poster format that displays both their results and claims based on the evidence from their analysis and their proposed study. During their research presentations, students participate in scientific argumentation by critiquing the claims of their peers and responding to the questions and comments of others, using a rubric designed by their teacher (5).

In developing the module, we adopted the following design principles, which could be applied to other programs:

Involving teachers, life scientists, and learning scientists as partners in curriculum design (6);

- Engaging students in topics that relate to their lives and interests (7);
- Developing analytical tasks that focus on the learning goals of the National Research Council science education framework (8) and the Next Generation Science Standards (<http://nextgenscience.org>);
- Providing instructional supports in the curriculum and database to guide student enactment of the research process (9);
- Designing the curriculum through an iterative process based on classroom implementation data (field notes, audio and video recordings, interviews, and survey items) and feedback from stakeholders;
- Providing teacher professional development workshops in which teachers

¹Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA. ²Learning Sciences and Institute for Science and Math Education, University of Washington, Seattle, WA 98195, USA. ³Department of Epidemiology, University of Washington, Seattle WA 98195, USA.

*IBI, Science Prize for Inquiry-Based Instruction; www.sciencemag.org/site/feature/data/prizes/inquiry/. †Present address: California Department of Public Health, Oakland, CA 94607, USA. ‡Corresponding author: mmunn@uw.edu

A

2. State your SPECIFIC hypothesis

2-1. You selected:
18.2. During your experimental smoking phase, did you experience any of the following regularly?
A "buzz"

2-2. Your SPECIFIC hypothesis for this question:
Smokers are more likely than people who didn't become smokers to have experienced a buzz during their experimental smoking period.

2-3. Your definition of "Exposed":
Feeling a buzz

2-4. Your definition of "Not-exposed":
NOT feeling a buzz

3. Identify answers corresponding to your exposure:

Exposure/Non-Exposure	Key Answer
Exposed Neither Not Exposed	a Yes
Exposed Neither Not Exposed	b No
Exposed Neither Not Exposed	c Don't know/not sure

B

Odds Ratio : 6.33 (Regular Smokers are 6.33 times more likely to have been exposed than non smokers.)

2 x 2 table

	Cases	Controls
Exposed	129	72
Not-Exposed	15	53

95% Confidence Interval : [3.33 , 12.02] (Your confidence interval is : [3.33, 12.02].)

Relationship between the Odds Ratio and Confidence Interval:

Database input and output pages. (A) Users select questions from the database, define specific hypothesis, "Exposed" and "Not-exposed;" select population, submit the query. (B) Interprets the result, and determine whether there is a causal relationship.

class for students for whom English is a second language, an Upward Bound seminar, advanced elective courses (e.g., genetics and biotechnology), and Advanced Placement and International Baccalaureate Biology, as well as community college courses. In the 2011–12 academic year, nearly 600 students participated in the project. Through feedback from teachers, classroom observations, and research studies conducted in the classrooms (10), we have learned several lessons regarding implementation of the module:

- Students often report that their interest in their investigation topics stems from their observations of smoking practices in their family, friends, and the media.

- Both students and teachers have limited prior experience and images of contemporary scientific practices other than the classical experimental design covered in traditional K–12 science curricula.

- Through their involvement in epidemiological research, students broaden their understanding of contemporary scientific research and methodologies, especially human subjects research.

- A study comparing student learning of scientific research and attitudes toward science after completing a genotyping experiment (3) and the database research described here. Students were somewhat more likely to rate genotyping as real science compared to their database research experience, despite associating more scientific tasks with the database experience than genotyping.

- Some students even expressed interest in conducting future case control studies focused on drug addiction in the future.

References and Notes

1. American Society of Addiction Medicine, "Public policy statement on nicotine dependence and tobacco" (ASAM, Chevy Chase, MD, 2010).
2. Centers for Disease Control and Prevention (CDC), *MMWR Morb. Mortal. Wkly. Rep.* **60**, 1207 (2011).
3. M. Munn et al., *Investigating the Effects of Genes and Environment on Smoking Behavior* (Univ. of Washington, Seattle, 2010); https://gsoutreach.gs.washington.edu/files/investigating_smoking_behavior_jan2011.pdf.
4. A. B. Hill, *Proc. R. Soc. Med.* **58**, 295 (1965).
5. P. Bell, in *Internet Environments for Science Education*, M. C. Linn, E. A. Davis, P. Bell, Eds. (Erlbaum, Mahwah, NJ, 2004), pp. 115–143.
6. National Research Council, *Learning Science in Informal Environments: People, Places, and Pursuits* (National Academies Press, Washington, DC, 2009).
7. P. Bell, M. C. Linn, *Int. J. Sci. Educ.* **22**, 797 (2000).
8. National Research Council, *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* (National Academies Press, Washington, DC, 2012).
9. M. C. Linn, E. A. Davis, P. Bell, *Internet Environments for Science Education* (Lawrence Erlbaum Associates, Mahwah, NJ, 2004).
10. H. Oura, K. Van Horne, A. W. Shouse, M. Munn, R. Knuth, in *The Future of Learning: Proceedings of the 10th International Conference of the Learning Sciences*, J. van Aalst, K. Thompson, M. J. Jacobson, P. Reimann, Eds., 2 to 6 July 2012, Sydney, Australia (ICLS and Univ. of Sydney, Sydney, 2012), vol. 2, pp. 41–48.

Acknowledgments: This project is funded by an NSF ITEST award, DRL-0929181. The predecessor project was supported through SEDAPA grant R25 DA013180 from National Institute on Drug Abuse (NIDA), NIH. The program content is solely the responsibility of the authors and does not necessarily represent the official views of the NSF, NIDA, or NIH. We wish to acknowledge project co-principal investigators P. Bell and J. Akey at the University of Washington; D. Nickerson and G. Jarvik, principal investigators of the predecessor project, the WISE development team in the Technology Enhanced Learning in Science center at the University of California, Berkeley, especially H. Terashima; many advising scientists, especially N. Weiss, UW Professor of Epidemiology; the many contributors to the curriculum, especially M. Brown; project evaluators R. Knuth and S. Levias; the many teachers who have implemented the curriculum in their classroom; and, of course, their students.

Supplementary Materials

www.sciencemag.org/cgi/content/full/341/6144/360/DC1

complete many of the student activities (see the first figure, right) and engage in discussions with lead teachers, scientists, education researchers, and each other; and

- Providing Web-based access to project components, implementation support, and ongoing technical support.

The module has been used in a wide variety of high school science courses, including introductory biology, a biotechnology

About the authors



(Left to right) **Maureen Munn** is the Director of Education Outreach in UW Genome Sciences, where she develops science education projects that integrate authentic genomics research into high school classrooms. **Hiroki Oura** is a Ph.D. candidate in the UW Institute for Science and Math Education focusing on curriculum, instructional, and technology design to develop students' scientific understanding and reasoning skills through authentic inquiry. **Mark Gallivan** is a M.P.H. graduate from the UW Department of Epidemiology and current Applied Epidemiology Fellow at the California Department of Public Health. **Katie Van Horne** is a graduate researcher at the UW Institute for Science and Math Education, focusing on how to engage all youth in contemporary scientific practices while taking into account their personal interests and identities. **Andrew Shouse** is the Associate Director of the UW Institute for Science and Math Education focusing on equitable science education in formal and informal settings and communication of research to policy and practice audiences.

PHOTO CREDITS: HANK CLARK, KATIE VAN HORNE, MIKA OTANI