International Journal of Science Education

ISSN: 0950-0693 (Print) 1464-5289 (Online) Journal homepage: http://www.tandfonline.com/loi/tsed20

Air Toxics Under the Big Sky: examining the effectiveness of authentic scientific research on high school students' science skills and interest

Tony J. Ward, Naomi Delaloye, Earle Raymond Adams, Desirae Ware, Diana Vanek, Randy Knuth, Carolyn Laurie Hester, Nancy Noel Marra & Andrij Holian

To cite this article: Tony J. Ward, Naomi Delaloye, Earle Raymond Adams, Desirae Ware, Diana Vanek, Randy Knuth, Carolyn Laurie Hester, Nancy Noel Marra & Andrij Holian (2016) Air Toxics Under the Big Sky: examining the effectiveness of authentic scientific research on high school students' science skills and interest, International Journal of Science Education, 38:6, 905-921, DOI: 10.1080/09500693.2016.1167984

To link to this article: http://dx.doi.org/10.1080/09500693.2016.1167984

Published online: 28 Apr 2016.

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION, 2016
VOL. 38, NO. 6, 905–921
http://dx.doi.org/10.1080/09500693.2016.1167984

Routledge
Taylor & Francis Group

Air Toxics Under the Big Sky: examining the effectiveness of authentic scientific research on high school students' science skills and interest

Tony J. Ward[a], Naomi Delaloye[a], Earle Raymond Adams[a], Desirae Ware[a], Diana Vanek[a], Randy Knuth[b], Carolyn Laurie Hester[a], Nancy Noel Marra[a] and Andrij Holian[a]

[a]Center for Environmental Health Sciences, University of Montana, Missoula, MT, USA; [b]Knuth Research Inc., Spokane, Washington, USA

ABSTRACT

Air Toxics Under the Big Sky is an environmental science outreach/education program that incorporates the Next Generation Science Standards (NGSS) 8 Practices with the goal of promoting knowledge and understanding of authentic scientific research in high school classrooms through air quality research. This research explored: (1) *how the program affects student understanding of scientific inquiry and research* and (2) *how the open-inquiry learning opportunities provided by the program increase student interest in science as a career path*. Treatment students received instruction related to air pollution (airborne particulate matter), associated health concerns, and training on how to operate air quality testing equipment. They then participated in a yearlong scientific research project in which they developed and tested hypotheses through research of their own design regarding the sources and concentrations of air pollution in their homes and communities. Results from an external evaluation revealed that treatment students developed a deeper understanding of scientific research than did comparison students, as measured by their ability to generate good hypotheses and research designs, and equally expressed an increased interest in pursuing a career in science. These results emphasize the value of and need for authentic science learning opportunities in the modern science classroom.

ARTICLE HISTORY
Received 23 June 2015
Accepted 15 March 2016

KEYWORDS
Environmental science; NGSS 8 Practices; open inquiry; secondary education; authentic science research

Introduction

Student interest in science

Recent trends have shown the US falling behind many developed countries in science education and student performance (The Organization for Economic Co-operation and Development, 2012). In addition, the rate of American students choosing careers in science, technology, engineering and mathematics (STEM) is lower compared to many other developed nations. Only a third of all first university degrees in the US being

CONTACT Tony J. Ward tony.ward@umontana.edu The Center for Environmental Health Sciences, University of Montana, Skaggs Building 32 Campus Drive, Missoula, MT 59812, USA

awarded are in science and engineering, compared to Japan and China where at least 50% of first-time degrees awarded are in the fields of science (National Science Board, 2014). Historically, more than half of freshman who declare STEM majors in the US leave these fields before graduating (Chen, 2013). Considering these trends, professionals in the field of education must continue to identify creative ways of promoting student interest in the fields of science within and outside the classroom.

This need to inspire our students in the field of science is benefitting from current trends in science education in America. With the development of the National Research Council's 'A Framework for K-12 Science Education' (NRC, 2012) and the *Next Generation Science Standards* (NGSS, 2013), teachers are being asked to reconsider their approach to teaching science. According to the NRC (2012), the current science education system is falling short as 'it is not organized systematically across multiple years of school, emphasizes discreet facts with a focus on breadth over depth, and does not provide engaging opportunities to experience how science is actually done' (p. 1). In response, the Framework emphasizes the 'integration of science and engineering practices to develop and use disciplinary core ideas (DCIs) and crosscutting concepts to explain phenomena and solve problems' (Krajcik, Codere, Dahsah, Bayer, & Mun, 2014, p. 158). Importantly, both the *Framework* and the NGSS emphasize learning the *process* of science along with the content and unifying learning across grades. This is not new phenomenon; Dewey (1910) lamented the lack of authentic scientific exploration in the science classroom more than a century ago and in the decades since inquiry-based learning has been a much-studied topic in science education (Barrow, 2006; Berg, Bergendahl, Lundberg, & Tibell, 2003; Minner, Levy, & Century, 2010).

While data still seem to be inconclusive on whether or not well-designed inquiry learning modules improve student learning of content over well-designed direct instruction modules, Cobern et al. (2010) point out that 'affective factors also play a role in learning, even if the focus is on science concepts; it may be that interest is sparked more naturally by inquiry, thus promoting positive attitudes toward science, which could lead to better performance' (p. 92). With this in mind, our project was designed to answer the following two questions:

(1) How does the *Air Toxics Under the Big Sky* program affect student understanding of the process of scientific inquiry and scientific research as compared to students in typical curricula?
(2) How do learning opportunities that incorporate open inquiry, such as those provided by the *Air Toxics Under the Big Sky* program, increase student interest in science as a career path?

Despite the recent focus on authentic science learning opportunities, and the potential positive impact on student interest in science, utilization of the resources and research available is still lacking. The NGSS 8 Practices are a model of inquiry-based learning that aim to address these shortcomings. As the NRC (2012) points out,

> the term 'inquiry', extensively referred to in previous standards documents, and has been interpreted over time in many different ways throughout the science education community. Our intent is to better specify what is meant by inquiry in science and the range of cognitive, social, and physical practices it requires. (p. 30)

For the same purpose of unifying the conversation around inquiry-based learning and in keeping with the forward trends in science education, our study focuses on the NGSS 8 Practices and how they serve as the basis for the open-inquiry model of learning provided by the *Air Toxics Under the Big Sky* program. Here we define open inquiry, as the NRC (2000) does: learning opportunities that require students to (1) engage in scientific questioning, (2) respond to questions using evidence, (3) formulate explanations from evidence, (4) connect explanations to scientific knowledge, and (5) communicate and justify explanations. In essence, open-inquiry learning 'simulates and reflects the type of research and experimental work that is performed by scientists' (Zion & Mendelovici, 2012).

The Air Toxics Under the Big Sky *program*

The *Air Toxics Under the Big Sky* program (described in Adams et al., 2008, 2009; Jones et al., 2007; Marra et al., 2011; Ward et al., 2008) was developed to create opportunities for students to participate in authentic science learning activities in secondary science education. The program seeks to provide students with learning opportunities that define progressive, student-centered, inquiry-based science education. As described by Anderson (2002) and Polman (2000), these involve active learning based on students' experiences and on their genuine interests. Polman (2000) emphasizes that since the time of Rousseau and Dewey, the themes inherent to constructivist science education continue to resurface decade after decade.

With the goal of engaging students in one of these meaningful learning experiences, students in the Air Toxics program participate in an open-inquiry research project related to issues of air quality and respiratory health in their homes and communities. While the general theme of the research project (air quality) is defined, the students have endless possibilities in what they study within the context of the project.

The Air Toxics program begins early in the school year when a researcher from the University of Montana visits each participating classroom and provides an overview of both indoor and outdoor air pollution issues, with an emphasis on $PM_{2.5}$ (airborne particulate matter $\leq$2.5 microns in aerodynamic diameter) in the western US and its impact on respiratory diseases such as asthma. The researcher remains an available resource to teachers and students throughout the school year based on the needs of the participants. Next, students work through three specially designed environmental health modules that address content about air pollution, health outcomes (e.g. asthma), and the use of air monitoring equipment. These three mandatory modules were designed to give students background information about air quality and environmental health in order to prepare students to identify their own testable questions and to conduct their own research.

Participating teachers choose when to use the modules and in which instructional units based on their regular curriculum and school schedules. Finally, classrooms are provided with an air sampling equipment and students are given comprehensive training in its operation. After completion of the three mandatory modules, students develop research questions, hypotheses, and research plans. Working in groups or independently, students use the air sampling equipment to collect data in support of their hypothesis-driven inquiry throughout the school year. Students collect air pollution data (i.e. $PM_{2.5}$) in

homes, schools, local businesses, and other locations relevant to their research in the community.

At the conclusion of the project, students analyze their data and communicate the results via posters or community reports, or by participating in the annual *Air Toxics Under the Big Sky* symposium on the University of Montana campus held each May (Vanek et al., 2011). At the symposium, students are asked questions by a small panel of experts in the air pollution/health field, who evaluate each project/presentation and determine the top three. From the outset, the Air Toxics program stressed that in addition to generating the research questions, hypotheses, and design, it is essential to make students responsible for communicating and justifying their findings (such as in the end-of-year symposium format). In a study focusing on classroom-based science inquiry from 1998 through 2007, Asay and Orgill (2010) found that less than a quarter of inquiry-based curricular initiatives emphasized *explaining and connecting* by students, and only about a tenth focused on having students *communicate and justify* their results to peers and experts. It is this full set of authentic scientific practices that motivates students to perform at higher levels, and provide a 'realistic' context in which an authentic inquiry process can unfold. In addition, preparing to respond to questions from an expert panel of judges encourages students to reflect on their research. Importantly, as presented in Table 1, each of the NGSS 8 Practices is covered as part of the *Air Toxics Under the Big Sky* program.

Methods

Teacher and student participants

Nine classrooms from five schools participated in the evaluation process. Teachers in the treatment group attended a curriculum workshop or worked individually with University of Montana staff to learn the program protocols and then agreed to fully implement the program curriculum in their designated classrooms. Full implementation was defined as: (1) offering the three mandatory lessons in their classrooms, (2) continued teacher support

Table 1. The NGSS 8 Practices and corresponding air toxics under the big sky activities.

Next Generation Science Standards 8 Practices	Corresponding air toxics under the big sky activity
1. Asking questions (for science) and defining problems (for engineering)	Students identify testable questions related to indoor/outdoor air quality in their community.
2. Developing and using models	Small-Scale Chemistry labs use 'Global Chambers' (i.e. petri dishes) to simulate and model a number of air pollution issues
3. Planning and carrying out investigations	Students design a research plan in order to investigate their testable question
4. Analyzing and interpreting data	Using data collected during their research, students analyze their results in order to draw conclusions
5. Using mathematics and computational thinking	During data analysis, students use graphs and basic statistical analyses to understand their results
6. Constructing explanations (for science) and designing solutions (for engineering)	Based on their data analysis, students draw conclusions about their air quality issue
7. Engaging in argument from evidence	Informal argumentation occurs when groups discuss results with one another as well as with the teacher.
8. Obtaining, evaluating, and communicating information	Students perform background research on their topic, integrate this into their project, and ultimately present their final work at end-of-year symposium or other community event

Table 2. Participants – treatment vs. control.

	Total number (*n*)	Gender	Age	Courses
Treatment	199	F: 55% M: 45%	18: 19% 17: 47% 16: 34%	Chemistry Human Anatomy Environmental Health
Control	180	F: 52% M: 48%	18: 32% 17: 43% 16: 25%	Biology Chemistry Human Anatomy Physics

of students in their data collection and research throughout the school year, and (3) helping students prepare for a public symposium presentation of their results. Teachers in the control group were recruited by teachers from treatment classrooms, and used a traditional science curriculum. All but one of the treatment teachers had prior experience with the *Air Toxics* curriculum.

All students participating in the evaluation process were in grades 11 and 12 ($n = 428$) and resided in northwestern Montana. The number of participating students who had both pre-survey and the content assessment exam available for analysis (methods described below) was 379 (180 comparison students, 199 treatment students). For more information on participating students, including gender and grade level, see Table 2.

Research design

Because it was not possible to randomly assign students into treatment or control groups, the establishment of experimental groups was based purely on the teachers' experience in the program described above and their willingness to use a modified curriculum. The control group students were from intact classrooms with teachers willing to volunteer as a comparison to the treatment groups. Both control and treatment groups were asked to complete a pre-survey at the beginning of the year, in addition to the content assessment exam at the end of the school year. Treatment students participating in the end-of-year symposium were also asked to complete a separate program evaluation form to determine the students' perspective of the program experience. The primary hypothesis being tested in this study was that the treatment group would perform better on the year-end content assessment compared to the control group students. In addition, the treatment group students would have a better concept of generating a research design and hypothesis. Details on all three surveys/assessments follow.

Pre-survey

The pre-survey of all participating students was conducted to determine if there were any variables that needed to be controlled in the ANOVA of the content assessment. Potential confounding variables (covariates) were assessed in the two experimental groups to determine their equivalence. The principle covariates used in this model included: (1) grade awarded in science the previous year, (2) grade awarded in math the previous year, (3) interest level in science content (Science Interest Scale), and (4) interest in performing environmental science tasks (Environmental Science Tasks Interest Scale). In addition, gender, school, grade level (11th or 12th), and the predicted number of science classes

in high school were examined as potential covariates. None of the latter were significantly different between groups.

Content assessment

Based on the learning objectives of the three mandatory lessons and the overall goals of the project, a 29 question exam, worth a total of 31 points, was constructed by the project team. Sixteen questions tested student knowledge of asthma and air pollution/particulate matter-related concepts. Ten questions tested general knowledge of research design, statistics, and data collection, while three questions had students evaluate whether or not a stated hypothesis was good or bad. Exam questions were primarily multiple choice with one important open-ended question to determine if the student could form a hypothesis and research design (see Appendix 1). The exam (administered in paper format) had 28 questions worth one point and one item (the opened-ended hypothesis/research question) worth three points. The content assessment was administered to students one week before the May symposium.

In the opened-ended hypothesis/research question, students were asked to generate a hypothesis on an open-ended item that they would like to investigate and then give a simple description of how they would test that hypothesis. Criteria for good and bad hypothesis/research design classifications were developed and a 'card sort' methodology was used to categorize responses into four piles (Figure 1). Three raters (all experts in environmental health related fields) were blinded to whether each response came from a student in the treatment or control group. The classification rules were:

- Good hypothesis: Had to be a statement of a relationship between variables, not a question; testable; and not nonsense (demonstrated understanding).
- Good research design: Had to be feasible; involve data collection; not violate ethics standards; and not nonsense (demonstrated understanding).

Inter-rater reliability, determined by Cronbach's alpha, was 0.70 (or 70%). Most of the disagreement between raters was specifically about the categorization of responses into the

	Bad Research Design	Good Research Design
Good Hypothesis	Pile One (2 Pts.)	Pile Three (3 Pts.)
Bad Hypothesis	Pile Zero (1 Pt.)	Pile Two (2 Pts.)

Figure 1. Scoring system for student-generated hypotheses and research design problems in the end of year content assessment. 'Pile' refers to student response categories.

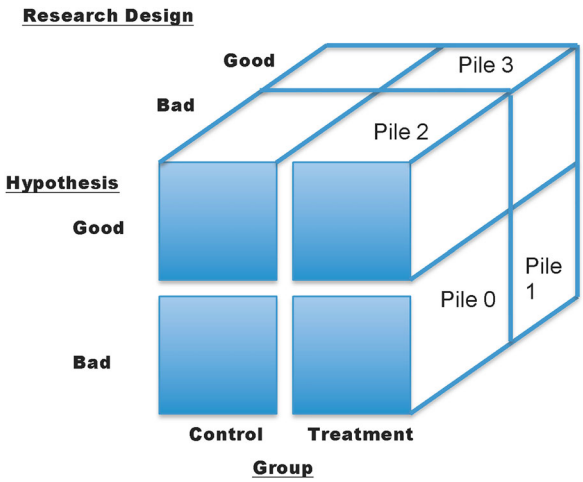


Figure 2. Model of three-way contingency table (2 × 2 × 2) for group vs. design. vs. hypothesis open-ended question rated by expert reviewers.

middle categories labeled Piles 1 and 2 (i.e. bad hypothesis/good research design or good hypothesis/bad research design). Agreement for placement into Pile 0 (bad hypothesis/bad research design) or Pile 3 (good hypothesis /good research design) was about 93%. A negotiation between raters was conducted which resulted in all responses being categorized into one of the four piles to enable further statistical analysis (three-level chi-square contingency table, see Figure 2).

Symposium survey

At the conclusion of the symposium, treatment students were additionally asked to complete a survey, which offered direct feedback from students on how the program affected their interest in science as content and as a potential career. In addition, their opinion about their experience with the program was assessed. The control group students did not attend the symposium, so there were no data for this group. This symposium survey was composed of Likert-scale rankings (ordinal-level data), in addition to open-ended responses related to their experience at the symposium and potential changes in their interest in science in general and as a potential career during the program. Responses from students to the open-ended questions were compiled and analyzed for repeating answers, which were then calculated as a percentage of total responses. A total of 448 students participated in the symposium surveys, including additional students who did not take the content exam or participate in the experimental part of this study.

Results

Differences between comparison and treatment groups in pre-survey variables

The analysis first explored the differences between comparison and treatment groups on the pre-survey covariates described above. First the reported grades from the previous year

were analyzed. The median grade in science for both comparison and treatment students from the previous year was an A-. In math, the median grade of comparison students was an A, while for treatment students it was an A-. In terms of this covariate, there was no significant difference between groups.

Additionally, no significant differences were found between comparison and treatment groups on the science interest scale. However, there were some minor differences in how students in each group rated their interest in performing activities that made up the environmental science task scale. Specifically, the treatment students tended to rate their interest slightly higher (significant at the $p \leq .01$ level). This covariate was controlled for in the final statistical model (ANCOVA). Control and treatment groups were also not significantly different with respect to the predicted number of years of science they would take in high school. More than 94% of students in both conditions expected to take four years of high school science. In general, the treatment and control groups were not statistically different in terms of past performance in science and math, and interest in science.

Performance differences on the content assessment

Figure 3 shows the difference between control and treatment groups with regard to performance on the content assessment exam, while Table 3 presents a breakdown for four of the content assessment exam subscales (asthma, $PM_{2.5}$, research, and hypotheses). Overall, treatment students scored higher on the overall content assessment than did students who did not participate in the program, with treatment students achieving 68.7% correct on the content assessment compared to only 45.9% correct for comparison students (Figure 3). Using ANCOVA (controlling for the influence of higher activity interest from the treatment group), the differences between treatment and comparison means were significant at $p < .001$. The following is a descriptive breakdown of performance in the different groups of questions.

Performance differences on asthma and $PM_{2.5}$ questions. On average, treatment students scored slightly higher on the asthma subscale, answering 3.7 out of 6 asthma questions

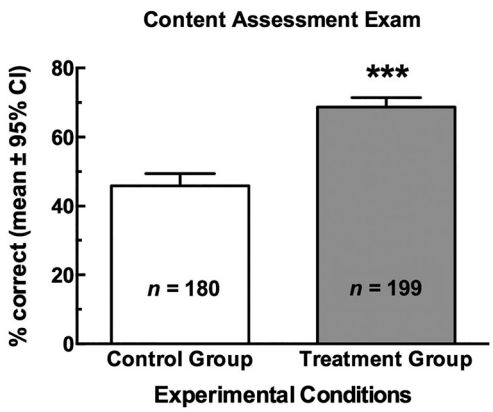


Figure 3. Average scores (% correct) of control and treatment groups in the end of year content assessment exam. Data expressed as mean ± 95% CI. Asterisks *** indicate $p < .001$ compared to the control group by ANCOVA with *post hoc* correction.

Table 3. Mean number of items correct on each content assessment subscale.

	Control			Treatment			Total		
Subscale	Mean	*N*	STD	Mean	*N*	STD	Mean	*N*	STD
Asthma (items 1–6; out of 6)	3.1	180	1.06	3.7	199	1.12	3.4	379	1.12
$PM_{2.5}$ (items 7–16; out of 10)	3.4	180	2.06	6.2	199	1.90	4.9	379	2.41
Research (items 17–25; out of 9)	5.3	180	1.68	6.7	199	1.47	6.0	379	1.70
Hypothesis (items 26–28; out of 3)	1.5	180	1.03	1.9	199	1.01	1.7	379	1.03

correctly while control students scored 3.1 out of 6 questions correct on the asthma subscale. Larger differences were noted in the $PM_{2.5}$ subscale with treatment students answering almost twice as many of 10 questions correctly as did control students (6.2 as compared to 3.4 items correct, respectively).

Performance differences on research design questions. Treatment students, on average, answered about 6.7 items out of 9 correctly whereas control students' average score was 5.3 items correct. A larger proportion of treatment students than comparison students were able to compute the mean of a list of numbers correctly. A higher percentage of treatment students also knew the appropriate statistical test to use in a hypothetical research scenario although the percentage correct for both groups was low (less than a third of all students).

Performance differences on hypothesis identification questions. Treatment students performed slightly better on the three items that asked students to evaluate whether or not a statement was a good or bad hypothesis (1.9 versus 1.5 out of 3 correct, on average).

Performance differences on the hypothesis and research design question. As Figure 4 shows, although some students in both treatment and comparison groups were awarded all three points, a considerably higher percentage of treatment students (62%)

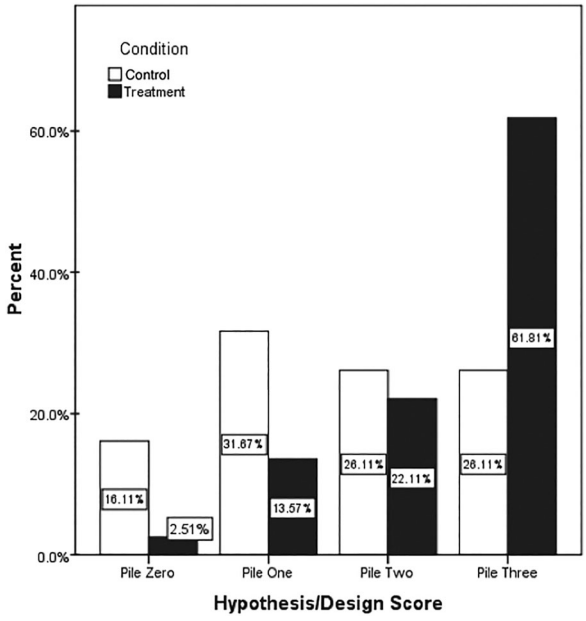


Figure 4. Differences in hypothesis/research design scores between treatment and control groups ($n = 379$).

received all three points on the hypothesis and research design question than did comparison students (26%).

About 16% of comparison students, compared to less than 3% of treatment students, received only one point on the hypothesis/research design item. More than half (58%) of comparison students received two points, while 36% of treatment students also received two points (Pile 1 and Pile 2). Moreover, considerably more comparison students gave nonsense answers or did not answer the question at all. Using a chi-square analysis, the difference in proportion between treatment and comparison groups was found to be significant, $\chi^2(\mathrm{DF} = 1, n = 379) = 42.82$ with correction for continuity, $p < .001$. Likewise, the difference in frequencies between group and hypothesis was also significant $\chi^2(\mathrm{DF} = 1, n = 379) = 12.45$ with correction for continuity, $p < .001$. Both analyses indicated that the good/bad design and hypothesis was not independent of group (control/treatment) (Figure 2). A simple examination of the numbers indicated that the treatment group was most associated with the 'good' design and 'good' hypothesis (Figure 4).

Symposium feedback results

At the conclusion of the *Air Toxics Under the Big Sky* symposium, 36% of students reported that they were more interested in science as a content area, and 24% reported an increased interest in a science career. Nearly a quarter of students listed their recognition that clean air was important. Students also communicated that learning about strategies to improve air quality was an important impact of participation in the *Air Toxics Under the Big Sky* program. Also listed was their increased knowledge about air pollution, improved public presentation skills, and, perhaps most importantly, an increased ability to conduct scientific research. In addition, an overwhelming percentage of students (85%) rated their experience in the symposium component of the program as either 'good' (59%) or 'excellent' (26%).

Discussion

The findings from this study highlight two very important points. First, students who have the opportunity to conduct authentic research of their own design, from the ground up, do indeed develop a deeper understanding of the processes of science compared to their counterparts who learned about these processes through lectures and labs commonly implemented within science classes. In many learning areas such as math, writing, music, art, and athletics, we emphasize the value of practice in the discipline. Yet we do not often apply the same model in our science classrooms. According to the NRC *Framework for K-12 Science Education* (2012), one of our current failures in science education is not 'provid[ing] students with engaging opportunities to experience how science is actually done' (p. 1).

With the development and broader teacher support of the NGSS 8 Practices, we see a much-needed movement toward more integration of authentic research into the classroom. Our results show that students performing research of their own design directly correlate with the skills addressed in the 8 Practices. *Air Toxics Under the Big Sky* students were better able to calculate mean numbers and identify correct statistical analyses, and were significantly better at writing a valid hypothesis and designing research. This

supports prior research that has demonstrated that the best way for students to become familiar with and skilled at science practices is by jumping in and doing them. Ebenezer, Kaya, and Ebenezer (2011) found that students developed proficient inquiry abilities in 7 of the 11 criteria they assessed, including defining a relevant scientific problem and systematically collecting and analyzing data with appropriate tools, while they participated in sustained research projects using technology to support data collection that focused on environmental issues in their communities.

The second significant finding is that student interest in science as a content area (36%) and potential career path (24%) increased after participating in the Air Toxics program. As discussed previously, to remain competitive in the global sphere of science education, we need more students pursuing science studies in their post-secondary education. More than one-third of students reported an increased interest in science after participation in the *Air Toxics Under the Big Sky* program, which demonstrates that learning experiences incorporating the NGSS 8 Practices (aka open-inquiry learning) are not only effective at improving student performance, but also boosting the overall interest in science.

These two findings support the idea that effective inquiry learning opportunities are needed in our schools. They address two necessary pieces to success: interest and skill. In one review of more than a decade of educational research (Potvin & Hasni, 2014), the authors found that the most significant predictors of student interest, attitude, and motivation in science learning were: (1) the teacher, (2) collaborative work in the classroom, and (3) meaningful learning 'where [Science and Technology] can be linked to reality' (p. 98). These were followed by hands-on, inquiry-based learning, lab activities, and learning opportunities that encouraged independent thinking. Programs like *Air Toxics Under the Big Sky* incorporate all of these into one well-rounded, year-long learning experience. And in doing so, this improves student's ability to do science, thereby increasing the chances that they may pursue science at a higher level.

In this study, no pre-assessment was administered at the beginning of the year, which would have helped track student improvement. Additionally, we were limited in the ability to track students as they went onto college, chose majors, graduated, and continued on a career path. It is only through a longer-term study such as this that we can definitively determine if programs like ours facilitate students ultimately working in a science profession. Regardless of this limitation, our research supports the growing body of research that indicates the value of and need for authentic research opportunities for secondary students.

We are interested in how our two findings are connected. Is there a direct relationship between students' understanding of the scientific process and their interest in a scientific career? Studies have already shown a connection between student self-concept, defined as a student's perceived ability to do well in a subject, and its crucial role in student success (Aschbacher, Ing, & Tsai, 2013; Lewis, Shaw, Heitz, & Webster, 2009). One study found that 'the strongest predictor of students' [Science/Engineering/Math] career interests in grades 7 to 9 and their aspiration trajectory over time were the questions about confidence as a science learner' (Aschbacher et al., 2013, p. 48). As science content becomes more rigorous over grade levels, students often begin to feel overwhelmed by the materials and, as a consequence, experience a lowered self-concept, eventually losing interest in pursuing a career in science. If programs like *Air Toxics Under the Big Sky* can provide a boost in confidence in a student's ability to do science, we hypothesize that this will also re-invigorate an interest in a scientific career.

Though this study demonstrates that students experience an increased understanding of the *process* of science, future studies should explore if the program equally increases comprehension and retention of content learning objectives over time. Students in both the comparison and treatment group were tested on knowledge of air quality and respiratory health at the end of the semester, though only treatment students received instruction in these specific areas. It would be valuable to investigate whether or not content knowledge is better retained over time by treatment students. As much as science skills are being emphasized in today's world of science education, content standards are still considered of greater importance. This being the case, examining content retention of standards-based knowledge should be examined in the future.

Conclusion

Overall, the *Air Toxics Under the Big Sky* program shows great promise for addressing current trends in science education. Through their research on air quality issues, students demonstrate an increased understanding of the processes of science, as well as an increased interest in a science career. Future studies should address how the program increases student understanding and retention of knowledge, over time, related to state and national content standards, as well as the relationship between increased understanding of scientific processes and increased interest in a science career. These findings are even more impactful when considering that participating schools are located in rural and underserved areas of western Montana, Idaho, and Alaska. A long-term goal of the *Air Toxics Under the Big Sky* program is to continue engaging students living in these disadvantaged areas, providing them with 'equal opportunities' to learn and explore the scientific process as it relates to air quality/respiratory health within their communities. Programs such as this have the potential to positively influence students' relationship with the field of science, and positively influence their future career choice.

Acknowledgements

We would like to thank Ray Hamilton for his assistance in statistical analysis.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the National Center for Research Resources and the Division of Program Coordination, Planning, and Strategic Initiatives of the National Institute of Health (NIH), through grant number [R25RR020432]. The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

Notes on contributors

Tony Ward is an Associate Professor at the University of Montana. He has both a BS (Sam Houston State University, Huntsville, Texas) and an MS (The University of Houston – Clear Lake, Houston,

Texas) in Environmental Science, and received his PhD in Chemistry from the University of Montana in 2001. In addition to conducting research within the Center for Environmental Health Sciences (CEHS), he is a core faculty member in the University of Montana's School of Public and Community Health Sciences, and is an adjunct faculty with the University of Montana's Department of Chemistry and Biochemistry.

Naomi Delaloye serves as the Education Coordinator at the Center for Environmental Health Sciences at the University of Montana, where she currently manages the *Air Toxics Under the Big Sky* program. She holds an MEd in Secondary Education and is a certified secondary science teacher with experience in classroom instruction, school curriculum development, and academic advising.

Earle Adams is a research assistant professor in the Chemistry Department at The University of Montana (UM) in Missoula. He has both a BS and an MS in Chemistry, and received his PhD in Analytical Chemistry. Earle was the recipient of two postdoctoral fellowships: one for five years at Yale University School of Medicine and the second for two years at Grinnell College in Iowa. Earle serves as the departmental coordinator for technology and tribal outreach and is the primary contact for the UM Research Experience for Undergraduate program, which supports active research participation by undergraduates in any of the areas of research funded by the National Science Foundation.

Desirae Ware is a research lab and project manager at the Center for Environmental Health Sciences at the University of Montana. She has a master's in Public Health and has had an active role in many projects with a wide range of focus from community awareness and health campaigns for people with disabilities to education outreach in rural communities through the Air Toxics program.

Diana Vanek is the outreach coordinator for the UM Center for Environmental Health Sciences and serves as co-investigator on the 'Environmental Health Science Education for Rural Youth' project funded by a Science Education Partnership Award (NIH Grant #1R25 RR020432). The Big Sky Model is drawn from one of the four main subprojects under this grant. Diana has a bachelor's and master's in anthropology, and has worked with advocacy, education, and workforce training programs serving disadvantaged populations in rural Montana communities. Her role at the UM CEHS involves promoting meaningful collaborations between university-based scientists and communities to enhance the public's role in identifying biomedical research priorities with the goal of improving environmental public health.

Dr Randy Knuth, president of Knuth Research, Inc. of Spokane, Washington, is the former science education evaluation specialist for *Air Toxics Under the Big Sky* project. He was actively involved in designing, implementing, and evaluating instructional systems for over two decades. Randy began his career in education as a high school teacher in Montana and completed graduate work in instructional technology and learning theory at Indiana University. This led to a research and development position at an educational research laboratory in Chicago where he developed instructional videos and Web sites, conducted research, and directed national projects. In addition to serving as the overall evaluator on the UM Environmental Health Science Education for Rural Youth program, Randy was involved with projects in the Northwest on geospatial science, biotechnology, health science, bullying prevention, and school-based mentoring.

Carolyn Hester serves as a Research Specialist III at the Center for Environmental Health Sciences at the University of Montana where she assists the Education Coordinator with the *Air Toxics Under the Big Sky* program and performs air quality research. She has a BS in Environmental Toxicology from the University of California at Davis and a secondary science teaching certificate from the University of Montana. Carolyn has experience teaching high school science and working in the environmental consulting field.

Nancy Noel Marra is the former Education Coordinator for the Center for Environmental Health Sciences (CEHS) at the University of Montana. Nancy has been involved in the field of education for over 25 years in many different capacities: classroom teacher, environmental education specialist,

gifted education specialist, Teacher Expectations and Student Achievement facilitator, differentiated instruction trainer, and student teacher supervisor. She has a master's degree in Education and was a recipient of the Presidential Award for Excellence in Science and Mathematics Teaching in 1996.

Dr Andrij Holian is the director of the UM Center for Environmental Health Sciences and is the principle investigator of the Environmental Health Science Education for Rural Youth project. Dr Holian received a PhD in Chemistry from Montana State University in 1975. He held positions at the University of Pennsylvania and Villanova University before joining the faculty of the University of Texas Health Science Center at Houston in 1984. During his tenure in Houston, he was also director of research for the Mickey Leland National Urban Air Toxics Research Center. In 2000 he arrived at the University of Montana to establish the Center for Environmental Health Sciences within the Department of Biomedical and Pharmaceutical Sciences. His current immunotoxicology research focuses on developing treatments to prevent lung inflammation and fibrosis resulting from exposure asbestos, silica, and nanoparticles.

References

Adams, E., Ward, T. J., Vanek, D., Marra, N., Hester, C., Knuth, R., … Boulafentis, J. (2009). The big sky inside: Measuring rural indoor air quality and its impact on the community. *The Science Teacher*, *76*(4), 40–45.

Adams, E., Ward, T., Vanek, D., Marra, N., Noonan, C., Smith, G., … Striebel, J. (2008). Air toxics under the big sky: A real-world investigation to engage high school science students. *Journal of Chemical Education*, *85*(22), 221–224.

Anderson, R. (2002). Reforming science teaching: What research says about inquiry. *Journal of Science Teacher Education*, *13*(1), 1–12.

Asay, L., & Orgill, M. (2010). Analysis of essential features of inquiry found in articles published in *The Science Teacher*, 1998–2007. *Journal of Science Teacher Education*, *21*(1), 57–79.

Aschbacher, P. R., Ing, M., & Tsai, S. M. (2013). Boosting student interest in science. *The Phi Delta Kappan*, *95*(2), 47–51.

Barrow, L. H. (2006). A brief history of inquiry: From Dewey to standards. *Journal of Science Teacher Education*, *17*(3), 265–278.

Berg, A. R., Bergendahl, C., Lundberg, B., & Tibell, L. (2003). Benefitting from an open-ended experiment? A comparison of attitudes to, and outcomes of, an expository versus an open-inquiry version of the same experiment. *International Journal of Science Education*, *25*(3), 351–372.

Chen, X. (2013). STEM attrition: College students' paths into and out of STEM fields statistical analysis report. Institute of Education Sciences, National Center for Education Statistics. Retrieved from http://nces.ed.gov/pubs2014/2014001rev.pdf

Cobern, W., Schuster, D., Adams, B., Applegate, B., Skjold, B., Undreiu, A., … Gobert, J. (2010). Experimental comparison of inquiry and direct instruction in science. *Research in Science & Technological Education*, *28*(1), 81–96.

Dewey, J. (1910). Science as subject-matter and as method. *Science*, *31*(787), 121–127.

Ebenezer, J., Kaya, O. N., & Ebenezer, D. L. (2011). Engaging students in environmental research projects: Perceptions of fluency with innovative technologies and levels of scientific inquiry abilities. *Journal of Research in Science Teaching*, *48*(1), 94–116.

Jones, D., Ward, T., Vanek, D., Marra, N., Noonan, C., Smith, G., & Adams, A. (2007). Air toxics under the big sky – a high school science teaching tool. *Science Education & Civic Engagement: An International Journal*, *1*(2), 51–55.

Krajcik, J., Codere, S., Dahsah, C., Bayer, R., & Mun, K. (2014). Planning instruction to meet the intent of the next generation science standards. *Journal of Science Teacher Education*, *25*(2), 157–175.

Lewis, S. E., Shaw, J. L., Heitz, J. O., & Webster, G. H. (2009). Attitude counts: Self-concept and success in general chemistry. *Journal of Chemical Education*, *86*(6), 744–749.

Marra, N., Vanek, D., Hester, C., Holian, A., Ward, T., Adams, E., & Knuth, R. (2011). Evolution of the air toxics under the big sky program. *Journal of Chemical Education*, *88*(4), 397–401.

Minner, D., Levy, A., & Century, J. (2010). Inquiry-based science instruction-what is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching*, *47*(4), 474–496.

National Research Council. (2000). *Inquiry and the national science education standards: A guide for teaching and learning*. Washington, DC: National Academies Press.

National Research Council. (2012). *A framework for k-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press.

National Science Board. (2014). *Science and Engineering Indicators 2014*. Retrieved from http://www.nsf.gov/statistics/seind14/index.cfm/chapter-1/c1h.htm

NGSS Lead States. (2013). *Appendix F of next generation science standards: For states, by states*. Washington, DC: The National Academies Press.

The Organization for Economic Co-operation and Development. (2012). *Program for International Student Assessment (PISA) results from 2012 survey*. Retrieved April 10, 2015, from http://www.oecd.org/pisa/keyfindings/PISA-2012-results-US.pdf

Polman, J. (2000). *Designing project-based science: Connecting learners through guided inquiry*. New York: Teachers College Press.

Potvin, P., & Hasni, A. (2014). Interest, motivation and attitude towards science and technology at K-12 levels: A systematic review of 12 years of educational research. *Studies in Science Education*, *50*(1), 85–129.

Vanek, D., Marra, N., Hester, C., Ware, D., Holian, A., Ward, T., … & Adams, E. (2011). The power of the symposium: Impacts from students' perspectives. *National Rural Education Association*, *32*(3), 22–28.

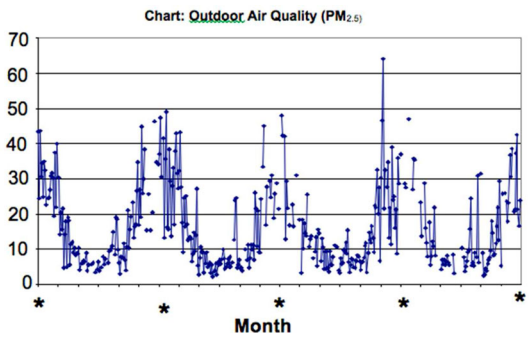
Ward, T. J., Vanek, D., Marra, N., Holian, A., Adams, E., Jones, D., & Knuth, R. (2008). The big sky model: A regional collaboration for participatory research on environmental health in the rural West. *Journal of Higher Education Outreach & Engagement*, *12*(3), 103–115.

Zion, M., & Mendelovici, R. (2012). Moving from structured to open inquiry: Challenges and limits. *Science Education International*, *23*(4), 383–399.

Appendix 1

Sample questions from the *Air Toxics Under the Big Sky* content assessment

(1) On the graph below of outdoor air quality ($PM_{2.5}$), which months of the year are most likely represented by each of the five asterisks (stars) along the bottom of the graph?
   (a) December, January, February
   (b) March, April, May
   (c) June, July, August
   (d) September, October, November



(2) A scientist is conducting research to find out whether noise levels increase the closer you get to bus stops. His data confirm the null hypothesis. What does this mean to the scientist?
(a) That the research hypothesis is valid.
(b) That the research hypothesis is true.
(c) That the null hypothesis is not valid.
(d) **That the research hypothesis is not valid.**

(3) What is the most important reason that the researcher should clean, maintain, and calibrate the sampler equipment?
(a) Keep the cost of maintenance down.
(b) **To be sure that data generated by the sampler are accurate.**
(c) Make sure it is ready for the next group of researchers.
(d) Keep the sampler from getting dirty.

(4) A group of students collected extensive air samples from two locations: 100 samples were taken near a bus stop and 100 were taken at the airport. The samplers were placed in the consistent locations at each location and data were collected at the same times of day. What statistical analysis could be used to see if there are differences between these two sets of data?
(a) Correlation Test
(b) **T-Test**
(c) Standard Deviation Test
(d) Mode Test

(5) For the table below, what is the mean of the data?
(a) 5.0
(b) 5.5
(c) **6.0**
(d) 6.5

Time (minutes)	Data ($PM_{2.5}$)
1	8.0
2	6.0
3	5.0
4	7.0
5	5.0
6	6.0
7	5.0
8	6.0
9	7.0
10	5.0

(6) While taking air samples with the equipment, Kelsey also kept notes about what she was observing. When she got back to the lab she noticed some spikes in her data. She checked her notes and remembered that some candles had been lit in the room. How should she treat the spikes in her data?
(a) Subtract out the spikes and use the average values instead

(b) **In the presentation use the documentation to explain the spike**
(c) Change the data to the average from the previous day
(d) Discard all of this day's data as it is no longer valid

(7) A solar eclipse happens when the moon blocks out the sun's light. Is this a good hypothesis?
(a) Good
(b) **Bad**

Explain why you think it is either 'good' or 'bad'.

This is a fact, not a hypothesis.

(8) Suppose you were asked to plan, carry out, and report on a research question dealing with air toxics. Please state the hypothesis you would like to investigate and how you would go about testing your hypothesis.