

## **Preliminary Results from the Thinking With Data Project: A Cross-Curricular Approach to the Development of Data Literacy**

Mark van 't Hooft, Annette Kratcoski, Dale Cook, Kent State University RCET  
Karen Swan, University of Illinois, Springfield  
Phil Vahey, Ken Rafanan, Louise Yarnall, SRI International

### **Abstract**

Data literacy is recognized as a critically important skill in today's society. While data literacy skills are included in disciplinary standards across the curriculum, they remain focused on de-contextualized data manipulation in Mathematics and are reduced to reading charts and graphs in most other subject areas. Nowhere is thinking with and about data really addressed. The Thinking with Data (TWD) project takes seriously the notion that data literacy is neither a single discipline nor a sub-discipline of Mathematics, but rather a skill that should be addressed in an interdisciplinary manner. This paper discusses what it means to address data literacy across the curriculum and shares the results of the field implementation of the TWD curriculum in two middle schools in Northeast Ohio.

### **Perspectives/Theoretical Framework**

Much has been written about the importance of understanding quantitative data in today's society (Briggs, 2002; Madison, 2002; Steen, 2001). Unfortunately, this realization has not translated into classroom practice. While there has been significant research on the teaching and learning of data analysis and probability (Lehrer & Schauble, 2001), and data analysis has been included in Mathematics education standards (NCTM, 2000), data analysis is still too often relegated to calculating measures of central tendency and reading simple graphs and tables, without aiming for true data literacy. Indeed, Rubin (2005, p. 22) writes that, "Numerical literacy' is woefully incomplete without 'data literacy,' yet we shortchange most students by leaving these topics out of the common series of math courses."

Although unfortunate, this situation is perhaps inevitable. Mathematics textbooks are already "a mile wide and an inch deep" (Schmidt, McKnight, Cogan, Jakwerth, & Houang, 1999), and data literacy takes significant time to develop. Data literacy includes the ability to formulate and answer data-based questions; use appropriate data, tools, and representations; interpret information from data; develop and evaluate data-based inferences and explanations; and use data to solve real problems and communicate their solutions. As such, true data literacy is neither a single discipline nor a sub-discipline of mathematics (Briggs, 2002; Madison, 2002; Scheaffer, 2001; Steen, 2001). This is most obvious in considering the role of the context of investigation: whereas in most mathematics "the context is part of the irrelevant detail...in data analysis, context provides meaning" (Cobb & Moore, 1997, pg. 801). We cannot expect this context for using the skills of data literacy to come solely from the mathematics classroom. True data literacy requires contributions from across the curriculum.

There is hope that contributions to data literacy may come from across the curriculum, as the importance of data literacy has been recognized across the disciplinary standards (see Table 1), and some of our prior research found significant student learning of key mathematical concepts in a three-week curricular unit that integrated mathematics and social studies through the use of

data (Vahey, Yarnall, Scan, Patton, & Zalles, 2006). By taking advantage of this confluence across the disciplines, we can allow students to experience true data literacy.

Table 1  
*Links to Data Literacy across Four Content Areas*

Data Literacy Requirement	Middle School Math Standards (NCTM)	Middle School SS Standards (NCSS)	Middle School Science Standards (NSES)	Middle School ELA Standards (NCTE)
Formulate and answer data-based questions	“Formulate questions, design studies, & collect data about a characteristic shared by two populations or different characteristics within one population.”	“Formulate historical questions, obtain historical data, question & identify gaps in data, & construct sound historical interpretations.”	“Identify questions that can be answered through scientific investigations. Develop the ability to refine & refocus broad & ill-defined questions.”	“Students conduct research on issues & interests by generating ideas & questions, & by posing problems.”
Use appropriate data, tools, and representations	“Select, create, & use appropriate graphical representations of data; discuss & understand the correspondence between data sets & their graphical representations.”	“Use appropriate geographic tools such as atlases, data bases, systems, charts, graphs, & maps to generate, manipulate, & interpret information.”	“Use appropriate tools & techniques to gather, analyze, & interpret data; the use of which, including mathematics, will be guided by the question asked & the investigations students design.”	“They gather, evaluate, & synthesize data from a variety of sources (e.g., print & non-print texts, artifacts, people) to communicate their discoveries in ways that suit their purpose & audience.”
Develop and evaluate data based inferences and explanations	“Use observations about differences between two or more samples to make conjectures about the populations; use conjectures to formulate new questions & studies to answer them.”	“Encourage increasingly abstract thought as learners use data & apply skills in analyzing human behavior in relation to its physical & cultural environments.”	“Students can learn to formulate questions, design & execute investigations, interpret data, use evidence to generate explanations, propose alternative explanations, & critique explanations & procedures.”	“Students use spoken, written, & visual language to accomplish their own purposes (e.g., for learning, enjoyment, persuasion, & the exchange of information).”

Alas, things are not so easy, as this requirement conflicts with the current organization and culture of our school system, which continues to treat the disciplines as separate, unrelated topics to be “covered” in 45-minute periods. The separation results in pedagogical cultures that miss opportunities to build on each other (Stevens, Wineburg, Herrenkohl, & Bell, 2005; Wineburg & Grossman, 2000). Most math classes, for example, limit students to approaching mathematics as exercises in number manipulation (Cobb & Bauersfeld, 1995), without thinking about real problems, nor pushing for evidence to back up claims (Kuhn, 1999). Unsurprisingly, students often fail to transfer and apply mathematical reasoning to understanding scientific content

(Akatugba & Wallace, 1999; Aldridge, 1994) or exploring societal problems. In social studies and English Language Arts, argumentation is rhetorical rather than quantitative (Stodolsky & Grossman, 1995). The divisions between these cultures interfere with students building data literacy.

## **The Thinking With Data Curriculum**

### **Data Literacy**

The Thinking With Data (TWD) project takes seriously the notion that robust learning is grounded in authentic contexts (Bransford, Brown, & Cocking, 2000; Brown, Collins, & Duguid, 1989). It is committed to engaging students in the investigation of a real societal problem using real world data across the disciplinary fields. Indeed, if the goal of data literacy is to develop students' abilities to use data to make informed decisions (Rubin, 2005; Steen, 2001), it only makes sense to so situate their learning. Therefore, the TWD unit (<http://www.rcet.org/twd/index.html>) was created as a series of four 2-week, integrated modules for cross-disciplinary implementation in 7<sup>th</sup>-grade social studies, mathematics, science, and English language arts (ELA) classes designed to develop students' deep understanding of data literacy across the curriculum. The social studies module sets the context -- fair allocation and use of water in the Tigris-Euphrates watershed as grounded in UN Resolution 51/229 (May, 1997) -- and introduces students to data literacy. The context was chosen because issues of water are among the most critical facing the world today (de Villiers, 2000), and because issues of "fairness" evoke investigations of real world data. In mathematics, this notion of "fairness" invokes the use of proportional reasoning. In science, students learn to use data to identify important patterns and relationships (Lehrer & Schauble, 2002), and to use data to develop coherent arguments (Kuhn & Udell, 2003) to address a larger social debate about water rights in the Middle East as well as in the United States. Finally, in ELA students use a variety of literacy skills, including thinking critically about concepts, claims, and arguments as well as reading, interpreting, evaluating, and synthesizing information (AASL, 1998) to communicate their findings, conclusions, and recommendations accurately and effectively.

Thus, the TWD curriculum focuses on the following aspects of data literacy:

- *Formulating and answering data-based questions.* Students formulate and answer questions based on their investigation of water availability data in social studies and their investigation of water quality data in science. These questions are deeply linked to core disciplinary content knowledge, and data analysis techniques learned in mathematics are applied to support student investigations in science and ELA.
- *Using appropriate data, tools, and representations.* In particular, the curricular materials focus on the central role of the creation of common measures as key to students' understanding what it means to use appropriate data and numerical representations. In our prior research (Vahey et al., 2006) we discovered that the analysis of many social studies contexts required a transformation of data, from raw values (such as the amount of water used by a set of countries, or their gross domestic products) to a measure that combined two quantities (Thompson & Thompson, 1992), such as a per capita measure. While the notion of transforming data to make it more meaningful is a key understanding in its own right, central to the work of creating common measures is the role of proportional reasoning. Proportionality is an essential middle-grades concept that can be used to make

sense of a variety of mathematical, scientific, and societal situations, and is a key element in thinking with data (Rubin, 2005). When embodied in authentic situations, proportionality entails multiple entry points for children's reasoning (Kaput & West, 1994; Lehrer, Strom, & Confrey, 2002), and is fundamental to productive growth in mathematical reasoning (Lamon, 1994). We also note that an understanding of proportionality is vital to understanding water quality data. For instance, the proportion of different macro-invertebrates in the form of a biotic index is a key indicator of stream health; similarly, the amount of phosphates in a sample of water from a stream is less important than its concentration in parts per million. In our focus on the central role of tools and visual representations, the TWD modules make use of readily-available, basic spreadsheet software to allow students to create representations of data sets and make comparisons between such data sets. These representations can be used to help students understand the value of unitizing measures, such as a per capita measure of water use.

- *Solving real problems and communicate their solutions.* Through the use of Web-based and paper materials, students are provided with the context needed to understand real social problems, such as how localities can address issues surrounding water availability and quality. Through the use of multimedia presentations students communicate their findings and solutions to their classmates. Future work will investigate the use of web-based media sharing tools for student communication of findings and solutions to a wider audience outside of the classroom.

As stated above, central to building students' data literacy is the overriding theme of *fairness*. In particular, the TWD unit focuses on the fairness of measures as they relate to fair allocation and fair comparison, as well as the fairness of data representations. Evaluating fairness is a productive activity for middle school students when engaged in data analysis (Hancock, Kaput, & Goldsmith, 1992; Lajoie, Jacobs, & Lavigne, 1995; Vahey et. al, 2000), and fairness is almost always deeply related with proportionality. The question of fair *allocation* of resources, such as water, almost always drives toward the notion of an allocation that is proportional, such as amount of water per the number of families or people in a region. Similarly, fair comparison, such as in comparing water quality or use, rests on the development of proportional measures. Fairness in *representation* – representing a complex group by a single individual, a complex situation by a single quantity, or a group of data points in a visual image – also hinges on proportionality. However, it additionally reveals another implied aspect of the utility of proportion: simple prediction by scaling. We would expect that a fair representation would allow, for example, doubling of some key aspect of the representation if the phenomenon under investigation is also doubled, and violation of this is often cited as a misleading use of representations (Tufte, 2001).

### **Preparation for Future Learning (PFL)**

While a detailed discussion of the Preparation for Future Learning framework and its impact on learning in TWD goes beyond the scope of this paper (and is the topic of another AERA paper), a brief explanation of this approach is necessary here in order to understand the TWD curriculum and materials for purposes of the research presented here. The PFL framework (Bransford & Schwartz, 1999) reverses the traditional lecture-and-apply process (Klahr & Nigam, 2004) and aligns with research that shows that students are more likely to learn when they have recognized the existence of a problem *before* being presented with a solution (Lehrer

& Schauble, 2002). In this framework students first *prepare* by investigating a set of problems that are designed to highlight the structure of an important concept. Instead of creating complete solutions, students come to understand the structure of the concept, and internalize key dimensions of the situation. Students then engage in a formal *learning* activity in which they are introduced to a standard solution, and which they then practice and apply in a variety of contexts. This instructional sequence is consistent with the conceptual change literature, which shows that students must first understand that there *is* a problem and then realize that their existing understandings are not adequate for creating a solution, before they are fully ready to learn scientific and/or mathematical concepts (Strike & Posner, 1992). PFL has been used in learning about statistics (Schwartz & Martin, 2004), the psychology of memory (Schwartz and Bransford, 1998), water quality (Barron et al., 1998), and ecosystems (Schwartz, Lin, Brophy, & Bransford, 1999).

### **Research Questions**

In order to determine the effectiveness of the TWD materials and their impact on student learning, we asked the following questions:

1. Are teachers able to effectively implement the cross-disciplinary TWD curriculum modules?
2. Do students who engage in the modules increase their understanding of cross-disciplinary data literacy at the end of the project versus the beginning of the project?
3. Do students who engage in the modules also increase their understanding of the required disciplinary content, particularly in Mathematics and Science?
4. Can we expect, based on teacher and principal input, that this program can be scaleable to a wide number of schools?

### **Methods**

#### **Procedures**

The TWD unit was pilot-tested with a group of 7<sup>th</sup> grade students (n = 42) from one middle school in Northeast Ohio during the 2007-2008 school year. Feedback from this pilot implementation resulted in substantial revisions of the individual modules, and all student and teacher materials were made available online (<http://www.rcet.org/twd/index.html>). Field implementation with 7<sup>th</sup> graders at two middle schools in Northeast Ohio followed during the fall semester of the 2008-2009 school year. A total of 114 students were taught using the TWD materials, and a total of 576 students participated in pre/post testing. Teachers received training by project staff related to the TWD materials and the PFL approach prior to implementing the individual modules. The modules were taught in sequence (social studies, math, science, and ELA) and as closely together as the school calendars allowed, with a maximum of one school week in between modules. Project staff provided teacher support as needed by phone, email, or face-to-face contact.

## Data Sources

During the field implementation, the following data were collected:

- *Pre/post data literacy assessments for the overall TWD unit.* The overall data literacy assessment was a problem-based exercise that tested whether students could use appropriate data, data representations, and tools to answer data-based questions and evaluate data-based inferences, explanations, and arguments. It consisted of five questions about a common problem that required interpretation of data from a table, synthesis of data between tables, manipulation of data to create common measures, interpretation of data presented in a narrative form, and an understanding of what data would be needed to answer questions more completely. The test instrument was designed by researchers on the TWD Project staff with input from the TWD advisory board to establish validity of the items. The test was administered immediately prior to the social studies module and following the ELA module (see Appendix 1).
- *Content assessments for the math and science modules.* The math and science content assessments were designed to assess students' ability to answer questions more directly related to the content standards in these disciplines. They were administered immediately prior to and following their respective modules.
- *Classroom observations.* At least one instance of every lesson by every teacher in every school was observed by an experienced RCET researcher (for a total of 81 observations), using a semi-structured observation format. Observers took detailed, running notes of key classroom behaviors and statements made by teachers and students, categorizing them on the fly using a small set of codes.
- *Student artifacts from all four modules.* Throughout the modules students created or modified discipline-relevant artifacts as they engaged in the activities. Artifacts include worksheets, reports, and essays created by students. Artifacts were collected for students from one of the schools; the other school did not have students keep their artifacts.
- *Student reflections about the TWD unit as a whole.* Students were asked to reflect on what they liked, didn't like, and learned in the TWD unit. Again, reflections were collected for students from one of the schools; the other school did not have students keep their artifacts.
- *Teacher and principal interviews conducted by the project's external evaluator.* The interviews were conducted by phone; teachers were interviewed immediately following the teaching of their respective modules, administrators were interviewed at the end of the entire project. The objective of the teacher interviews was to get teacher's perceptions of how difficult it was to teach the modules, what impact they had on their students' learning and behaviors, and what recommendations they had for further dissemination of the modules. The administrators were interviewed to determine their reaction to the project and to obtain any ideas for further dissemination of the modules. Interview protocols can be found in Appendix 2.

## Analysis and Results

### Effective Teacher Implementation

Effective teacher implementation (research question 1) is crucial if the TWD unit is to yield positive learning outcomes. We used two data sources to analyze this question: teacher interviews (conducted by our external evaluator), and classroom observations. Interview transcripts and observation field notes were qualitatively analyzed to tease out patterns and themes. Teacher interviews indicated that the modules could be successfully implemented with the appropriate administrative support and professional development. The most difficult module was the first one, social studies. There were two reasons for this: First, the content of the module did not perfectly align with the 7<sup>th</sup> grade social studies curriculum, and was considered more of a review of 6<sup>th</sup> grade material (this was due to a change in standards from the time the project was funded to the time of our field test). Fortunately, the social studies component is implemented first in the TWD progression, so teachers could implement it early in the year when such a review was still appropriate. Second, while consistent with PFL and essential for setting up student learning in the math module, the students and teachers did not have closure at the end of the social studies module concerning equitable distribution of water among the countries of the Tigris/Euphrates watershed (closure came from the analyses conducted in math class). Both social studies teachers expressed some frustration about this lack of closure. Teachers in math and science did not have these concerns, as these modules were better aligned with their 7<sup>th</sup> grade curriculum, and they both had adequate closure with respect to content goals. ELA was an opportunity for students to show their integrated understanding of cross-disciplinary data literacy, and according to the teachers this “went very well.” All math, science, and ELA teachers reported being very satisfied with the unit. Classroom observations also indicated that all teachers, including social studies teachers, were able to implement the materials at some level of competence. Teachers were able to “cover” most of the expected materials each day, students were engaged during the lessons, there were no obstacles that prevented completion of key components of the materials, and all teachers completed all lessons in their module in the expected time. However, as will be discussed below, there was a high level of variability in implementation, and the funds from this supplement will be used to investigate this variability in more detail.

### Understanding of Cross-Disciplinary Data Literacy

Regarding data literacy across the entire unit (research question 2), analysis of pre/post test data indicated that, on average, students who were exposed to the TWD curriculum ( $n = 114$ ) had a gain score that was three points higher on a 15-point test than comparable students who did not engage with the TWD materials ( $n = 462$ ). This difference was statistically significant:  $t(156.273) = 10.750, p < .001, d = 1.24$  (very large effect).

When analyzing the scores for each school individually we found differences between them in student performance (Table 2). TWD students at School 2 seemed to have learned more than those at School 1 with mean gain scores at 3.69 and 2.03, respectively. At School 2, the mean difference in gain scores between the TWD group and their 7<sup>th</sup> grade classmates was 3.13, as compared to a mean difference of 2.05 at School 1. However, effect sizes for both schools were large (1.27 and 1.00 respectively).

Table 2  
*TWD IMD Pre/Post Test Demographics*

School	# of students		Gain Scores		Gain Score Difference
	TWD	No TWD	TWD	No TWD	
1	29	174	2.03	-0.02	2.05
2	85	288	3.69	0.56	3.13
Total	114	462			

MANOVA was used to determine gain scores from pre to post test for each question individually, as well as to calculate effect sizes. For the overall MANOVA, the Box *M* test of equality of covariance matrices yielded an *F*-statistic of 6.172,  $p < .000$  (i.e. equality of variance was violated). Hence, the MANOVA result reported here is for the corrected model. The overall multivariate test yielded an  $F(5, 570) = 32.92$ ,  $p < .000$ , with a partial eta-squared = .308, meaning that approximately 31% of the difference in total gain scores between the experimental and control groups can be explained by the TWD intervention.

Similarly, homogeneity of variance for each question separately was violated (Table 3).

Table 3  
*Levene's Test of Equality of Error Variances(a) for Gain Scores by Question Analysis*

Gain Score/Question	F	df1	df2	Sig.
diff_pre_post_1	19.164	1	574	.000
diff_pre_post_2	7.211	1	574	.007
diff_pre_post_3	4.247	1	574	.040
diff_pre_post_4	83.852	1	574	.000
diff_pre_post_5	44.518	1	574	.000

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+TWDunit

Between-subject tests are shown in Table 4 below. Note that the *F*-statistic for each of the tests is statistically significant at  $p < .000$ , and that the partial eta square statistics indicate approximately how much of the variability in gain scores between the experimental and control groups can be explained by each question individually (3.9 %, 9.5 %, 3.0 %, 14.4 %, and 4.9 % respectively). The data indicate that Question 4 explains the largest proportion of difference.

Table 4  
*Tests of Between-Subjects Effects for Gain Scores by Question Analysis*

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	diff_pre_post_1	22.662(b)	1	22.662	23.249	.000	.039
	diff_pre_post_2	41.462(c)	1	41.462	60.470	.000	.095
	diff_pre_post_3	25.121(d)	1	25.121	17.727	.000	.030
	diff_pre_post_4	87.787(e)	1	87.787	96.681	.000	.144
	diff_pre_post_5	10.383(f)	1	10.383	29.866	.000	.049

In sum, when considering questions individually, student scores improved the most on those items that required higher order thinking skills such as data interpretation across multiple tables and calculation of proportional data, providing support for the argument that the TWD materials aid in the development of data literacy skills among middle school students.

The next logical step was to analyze the pre-post test results by question and by school. At School 1, three 7<sup>th</sup> grade teams participated in the study (n = 203). Data was used for 29 students who were in the experimental group and 174 students who were divided over two control groups. The Box M test of equality of covariance matrices for test data at School 1 yielded  $F = 1.587$ ,  $p = .069$  (i.e. equality of variance was NOT violated). The overall multivariate test yielded  $F(5, 197) = 6.465$ ,  $p < .000$ , with a partial eta-squared = .141. Approximately 14% of the difference between total gain scores of the experimental and control groups at School 11 can be explained by the TWD intervention.

Homogeneity of variance for each question separately was violated for only one out of the five test questions, i.e. Question 4 (Table 5).

Table 5

*Levene's Test of Equality of Error Variances(a) for Gain Scores by Question by School Analysis (School 1)*

Gain Score/Question	F	df1	df2	Sig.
diff_pre_post_1	.566	1	201	.453
diff_pre_post_2	1.440	1	201	.232
diff_pre_post_3	2.129	1	201	.146
diff_pre_post_4	4.736	1	201	.031
diff_pre_post_5	1.092	1	201	.297

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a Design: Intercept+TWDunit

Between-subject tests are shown in Table 6. Note that the  $F$ -statistic for each of the tests is statistically significant at  $p < .05$ , except for Question 3 ( $F = 2.007$ ,  $p = .158$ ). The partial eta square statistics indicate approximately how much of the variability in gain scores between the experimental and control groups can be explained by each question individually (3.5 %, 9.0 %, 1.0 %, 3.8 %, and 2.8 % respectively). The data indicate that Question 2 explains the largest proportion of difference between the experimental and control groups at School 1.

Table 6

*Tests of Between-Subjects Effects for Gain Scores by Question by School Analysis (School 1)*

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	diff_pre_post_1	6.214(b)	1	6.214	7.378	.007	.035
	diff_pre_post_2	12.021(c)	1	12.021	19.992	.000	.090
	diff_pre_post_3	2.667(d)	1	2.667	2.007	.158	.010
	diff_pre_post_4	5.387(e)	1	5.387	7.930	.005	.038
	diff_pre_post_5	1.971(f)	1	1.971	5.754	.017	.028

At School 2, four 7<sup>th</sup> grade teams participated in the study (n = 373). Data was used for 85 students who were in the experimental group and 288 students who were divided over three control groups. The Box *M* test of equality of covariance matrices for test data at School 2 yielded an *F*-statistic of 5.775,  $p < .000$  (i.e. equality of variance was violated). Hence, the MANOVA result reported here is for the corrected model. The overall multivariate test yielded  $F(5, 367) = 25.204$ ,  $p < .000$ , with a partial eta-squared = .256. Hence, approximately 26% of the difference in total gain scores between the experimental and control groups at Kimpton Middle School can be explained by the TWD intervention.

Homogeneity of variance for each question separately was violated for four out of the five test questions (Table 7).

Table 7

*Levene's Test of Equality of Error Variances(a) for Gain Scores by Question by School Analysis (School 2)*

Gain Score/Question	F	df1	df2	Sig.
diff_pre_post_1	15.170	1	371	.000
diff_pre_post_2	3.310	1	371	.070
diff_pre_post_3	1.795	1	371	.181
diff_pre_post_4	60.021	1	371	.000
diff_pre_post_5	52.515	1	371	.000

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a Design: Intercept+TWDunit

Between-subject tests are shown in Table 8. The *F*-statistic for each of the tests is statistically significant at  $p < .000$ , and the partial eta square statistics indicate that the amount of variability in gain scores between the experimental and control groups explained by each question individually is 3.5 %, 9.1 %, 4.0 %, 18.4 %, and 5.6 % respectively. The data indicate that Question 4 explains the largest proportion of difference.

Table 8

*Tests of Between-Subjects Effects for Gain Scores by Question by School Analysis (School 2)*

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	diff_pre_post_1	14.178(b)	1	14.178	13.647	.000	.035
	diff_pre_post_2	27.167(c)	1	27.167	37.257	.000	.091
	diff_pre_post_3	22.546(d)	1	22.546	15.366	.000	.040
	diff_pre_post_4	79.029(e)	1	79.029	83.746	.000	.184
	diff_pre_post_5	7.742(f)	1	7.742	22.152	.000	.056

In sum, the results of the pre-post test analysis indicate that the TWD unit had an impact on students' data literacy skills as tested by the pre/post test. The mean difference in gain scores between the experimental and control groups was about 3 points on a test with a maximum score of about 15 points, or roughly a 20% difference. That this is a large difference is indicated by an associated effect size of 1.24. When looking at each school individually, there is a difference in overall student performance. In general, students at School 2 seemed to have fared better than those at School 1, both when looking at overall gains as well as gain scores by question.

When the focus of the analysis is on the individual questions, the largest differences in gain scores occurred on question 4 (14.4%), and 2 (9.5%) for the full sample. Both of these questions required higher order thinking skills such as data interpretation across multiple tables and calculation of proportional data, thus providing support for the argument that the Thinking With Data materials aid in the development of data literacy skills among middle school students. Test results for School 2 (and to a lesser extent for School 1) seem to uphold this argument as well. At School 2, students improved most on questions 4 and 2; at School 1, they mostly improved on question 2.

Finally, the results of this pre/post test need to be interpreted with caution. For one, there was a substantial difference in size between the experimental and control groups (144 and 462 students respectively). In addition, homogeneity of covariance and variance were often violated, opening the door for statistical imprecision or error. A good example of this can be found when comparing overall MANOVAS with their respective between-subject tests, which shows a substantial amount of overlap of variance explained by the individual questions on the data literacy pre-post tests. Finally, even though the overall effect sizes are large, when studying differences in gain scores in more detail, it appears that the test is an example of many small differences adding up to a larger one.

### **Understanding of Disciplinary Content**

Pre/post testing in math and science showed statistically significant gains in learning disciplinary content (research question 3). Due to inconsistencies in test administration, only two items could be scored on the Math test for School 1; both items showed statistically significant gains ( $Z = 3.16$  and  $Z = 4.70$  respectively). Students in School 2 showed statistically significant gains across the entire math test,  $t(24) = 4.899$ ,  $p < .001$ ,  $d = .56$  (medium effect). In science, students in both schools showed statistically significant gains too, with  $t(84) = 12.665$ ,  $p < .001$ ,  $d = 1.36$  (very large effect) for School 1; and  $t(27) = 4.441$ ,  $p < .001$ ,  $d = .83$  (large effect) for School 2. Further analysis of test data is currently in progress.

### **Scalability**

Finally, teachers and principals indicated in their interviews that the TWD unit is scalable (question 4). With the exception of social studies, teachers felt that the TWD unit fit into the 7<sup>th</sup> grade curriculum. There were concerns about the tensions created when modules are taught using PFL, especially at the end of the social studies module. However, in our interactions with the social studies teachers following their implementation of the module, we found that these tensions were alleviated somewhat by the knowledge that students did have closure through participation in the remaining disciplines, and that students' ELA assignments were of high quality and showed that they could integrate their social studies knowledge with what they learned in other disciplines. A small number of teachers were concerned that some of the materials were too advanced for their students, and these concerns will likely be addressed in our final set of materials by including alternative options for lower-achieving students. Finally, teachers expressed the importance of staff development, and working as a team of teachers so each knew what the other three were doing in their modules. The administrators agreed with the teachers about the usefulness of the TWD unit, and stressed the importance of alignment with the

content standards, as well as the importance of teachers working as a team, with each teacher being responsible for only one subject area.

## Conclusions

Project findings up to this point are significant in the following ways. First, we created and investigated the effectiveness of a curriculum based on the PFL framework of teaching and learning. We found that *preparing* students for data literacy learning can occur in one curricular context with the *learning* activity occurring in another, strengthening the plausibility of claims that PFL uncovers a general mechanism of transfer (Bransford & Schwartz, 1999). We also found that subsequent application and communication of learning may amplify learning that happens when using PFL, extending the framework (a detailed discussion related to the use of PFL for TWD is the subject of another AERA paper and goes beyond the scope of this one). In other words, our findings suggest that projects such as TWD work well when learning activities are integrated across multiple disciplines. Furthermore, this study provides a theoretically and empirically grounded basis for increasing the use of real-world data and developing students' data literacy across the curriculum, providing a benchmark against which other such efforts can be measured.

Second, with its materials having been tested in an authentic school setting, the TWD Project provides a scientific basis for conducting school-based data literacy activities that cut across disciplines; provides a set of assessment tools that can be used to investigate students' formative reasoning as they engage in cross-disciplinary data literacy; and provides a collection of materials and related data sets that can be used by others in the implementation of cross-disciplinary data literacy.

In sum, our original TWD project sought to create a cross-curricular unit to develop data literacy skills in middle school students and focused on materials development and teacher training. Preliminary findings suggest the effectiveness of the TWD approach. Future work will focus on more closely analyzing our existing data to investigate the specific conditions that allow for our PFL-based approach to be effective, and on investigating how students and teachers make cross-disciplinary links when using the materials. In short, we know that the TWD materials were effective; we need to investigate *why* and *how* they worked as well as they did. This work will especially focus on student work and reflections.

## Acknowledgements

The research reported on in this paper was funded by the National Science Foundation under Grant NO. NSF ESI-0628122. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

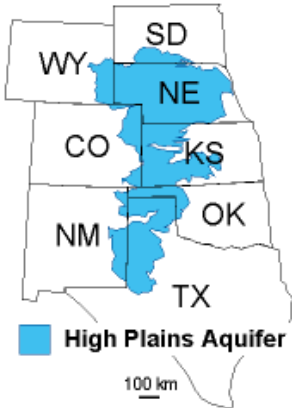
## References

Akatugba, A. H., & Wallace, J. (1999). Mathematical dimensions of students' use of proportional reasoning in high school physics. *School Science and Mathematics*, 99(1), 31-39.

- Aldridge, B. (1994). The role of mathematics in science education. *Focus on Learning Problems in Mathematics*, 16(5), 36-44.
- American Association of School Librarians (AASL). (1998). *Information literacy standards for student learning: Standards and indicators*. Retrieved from [http://www.ala.org/aasl/ip\\_nine.html](http://www.ala.org/aasl/ip_nine.html)
- Bransford, J. D., Brown, A. L., & Cocking. R. R. (2000). *How People Learn*. Washington, DC: National Academy Press.
- Bransford, J. D., & Schwartz, D. (1999). Rethinking transfer: A simple proposal with multiple implications. In A. Iran-Nejad & P. D. Pearson (Eds.). *Review of Research in Education: Vol. 24* (pp. 61-100). Washington, DC: American Educational Research Association.
- Briggs, W. (2002). Quantitative literacy and SIAM. *SIAM News* 35(2). Retrieved from [http://www.Math.cudenver.edu/~wbriggs/qr/siam\\_news.html](http://www.Math.cudenver.edu/~wbriggs/qr/siam_news.html)
- Brown, J. S., Collins, A. & Duguid, A. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32-42.
- Cobb, G. & Moore, D. (1997). Mathematics, statistics, and teaching. *American Mathematical Monthly*. 104, 801- 823.
- Cobb, P., & Bauersfeld, H. (Eds.). (1995). *Emergence of mathematical meaning: Interaction in classroom cultures*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cognition and Technology Group at Vanderbilt. (1997). *The Jasper project: Lessons in curriculum, instruction, assessment, and professional development*. Hillsdale, NJ: Erlbaum.
- Cognition and Technology Group at Vanderbilt. (1992). The Jasper series as an example of anchored instruction: Theory, program description and assessment data. *Educational Psychologist*, 27, 291-315.
- de Villiers, M.(2000). *Water: The fate of our most precious resource*. New York: Houghton Mifflin Company.
- Hancock, C., Kaput, J., & Goldsmith, L. (1992). Authentic inquiry with data: critical barriers to classroom implementation. *Educational Psychologist*, 27(3), 337-364.
- Kaput, J., & West, M. (1994). Missing-vale proportional reasoning problems: Factors affecting informal reasoning patterns. In G. Harel & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics* (pp. 235-287). Albany: State University of New York Press.
- Klahr, D., & Nigram, M. (2004). The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. *Psychological Science*, 15, 661–667.
- Kuhn, D. (1999). A developmental model of critical thinking. *Educational Researcher*, 28(2), 16-26, 46.
- Kuhn, D., & Udell, W. (2003). The development of argument skills. *Child Development*, 74(5), 1245–1260.
- Lajoie, S., Jacobs, V., & Lavigne, N. (1995). Empowering students in the use of statistics. *Journal of Mathematical Behavior*, 14, 401-425.
- Lamon, S. (1994) Ratio and proportion: Cognitive foundations in unitizing and norming. In J. Confrey & G. Harel (Eds.), *The development of multiplicative reasoning in the learning of mathematics*. (Research in Mathematics Education Series, pp. 89-122). Albany, NY: State University of New York Press.
- Lehrer, R., & Schauble, L. (2002). *Investigating real data in the classroom: Expanding children's understanding of math and science*. New York: Teachers College Press.

- Lehrer, R., Strom, D., & Confrey, J. (2002). Grounding metaphors and inscriptional resonance: Children's emerging understanding of mathematical similarity. *Cognition and Instruction*, 20(3), 359-398.
- Madison, B. L. (2002, February). Educating for numeracy: A challenging responsibility. *Notices of the American Mathematical Society* 49(2). Retrieved from [www.ams.org/notices/200202/commentary.pdf](http://www.ams.org/notices/200202/commentary.pdf)
- National Council of Teachers of Mathematics (NCTM). (2000). *Standards 2000*. Reston, VA: National Council of Teachers of Mathematics.
- Rubin, A. (2005, Spring). Math that matters. *Threshold*, 22-31. Retrieved from <http://www.ciconline.org/NR/rdonlyres/eqegvj373bnr4m5wektaq33yqm6f3hyk77bjca24ywlyyvqfqvaltyeo7qzrejvgtwz4nlu5ncpeewdrzwcmdzpksh/T-Spr05-MaththatMatters.pdf>
- Schmidt, W. H., McKnight, C., Cogan, L. S., Jakwerth, P. M., & Houang, R. T. (1999). *Facing the consequences: Using TIMSS for a closer look at U.S. mathematics and science education*. Dordrecht, Netherlands: Kluwer.
- Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction*, 16(4), 475-522.
- Schwartz, D. L., Lin, X. D., Brophy, S., & Bransford, J. D. (1999). Towards the development of flexibly adaptive instructional design. In C. Reigeluth (Ed.), *Instructional design theories and models* (Volume II, pp. 183-213). Mahwah, NJ: Lawrence Erlbaum Associates.
- Schwartz, D. L., & Martin T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction *Cognition and Instruction*, 22(2), 129-184.
- Steen, L.A. (2001). *Mathematics and democracy: The case for quantitative literacy*. USA: National Council on Education and the Disciplines.
- Stevens, R., Wineburg, S., Herrenkohl, L. R., & Bell, P. (2005). The comparative understanding of school subjects: Past, present, and future research agenda. *Review of Educational Research*, 75(2), 125-157.
- Stodolsky, S. S., & Grossman, P. L. (1995). The impact of subject matter on curricular activity: An analysis of five academic subjects. *American Educational Research Journal*, 32, 227-249.
- Strike, K. A., & Posner, G. J. (1992). A revisionist theory of conceptual change. In R. A. Duschl & R. J. Hamilton (Eds.), *Philosophy of science, cognitive psychology and educational theory and practice* (pp. 147-176). Albany, NY: State University of New York Press.
- Thompson, P. W., & Thompson, A. G. (1992, April). Images of rate. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.
- Tufte, E. R. (2001). *The visual display of quantitative information (2<sup>nd</sup> edition)*. Chester, CT: Graphics Press.
- United Nations. (May, 1997). *UN Convention on the Law of the Non-navigational Uses of International Watercourses* (Resolution 51/229). New York: General Assembly of the United Nations.
- Vahey, P. J., Yarnall, L. G., Scan, K. P., Patton, C., & Zalles, D. R. (2006). Mathematizing middle school: Results from a cross-disciplinary study of data literacy. Paper presented at the 2006 Annual Meeting of the American Educational Research Association, San Francisco, CA.
- Wineburg, S., & Grossman, P. (Eds.). (2000). *Interdisciplinary curriculum: Challenges to implementation*. New York: Teachers College Press.

## Appendix 1: Data Literacy Pre-Post Test



The High Plains (or Ogallala) Aquifer is an underground reservoir of water. It contains about as much water as Lake Huron. Most of the aquifer’s water was deposited during the Ice Ages. The High Plains Aquifer lies under 174,000 square miles of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, and Wyoming. The depth of the aquifer ranges from 0 feet at its boundaries to more than 1,000 feet in west-central Nebraska.

The High Plains Aquifer is the most important source of water in the High Plains region. It provides nearly all the water for residential, industrial, and agricultural use. Irrigated agriculture forms the base of the regional economy. It accounts for 94% of the use of water drawn from the aquifer. It also supports nearly one-fifth of the wheat, corn, cotton, and cattle produced in the United States.

In the 1940s, farmers began pumping large amounts of water for irrigation from the High Plains Aquifer. Farmers were taking more water from the aquifer than was naturally replaced by water seeping into it from the surface. By 1980, water levels had dropped by over 100 feet in parts of Texas, Oklahoma, and Kansas.

People decided something had to be done. Government programs encouraged **better irrigation and more efficient agricultural practices**. The leaders of this movement believe it has been successful. They say that the decline in water levels has slowed since 1980 when changes were made (Table 1).

	Percentage of aquifer water	1950-1980 water level change in ft	1980-1995 water level change in ft	1996 water level change in ft
Colorado	3.7	-4.2	-5.9	+0.15
Kansas	9.9	-9.9	-8.0	+0.23
Nebraska	65.5	0.0	+2.3	+0.58
New Mexico	1.5	-9.8	-4.0	-0.65
Oklahoma	3.4	-11.3	-2.7	+0.70
South Dakota	1.8	0.0	+3.1	+0.82
Texas	12.0	-33.7	-7.4	-0.98
Wyoming	2.2	0.0	-1.5	-0.24
<b>High Plains Aquifer</b>	<b>100</b>	<b>-9.9</b>	<b>-0.33</b>	<b>+0.08</b>

Table 1: Changes in Water Levels in the High Plains Aquifer, 1950-1997

Others argue that **above average rain** from 1980 to 1996 is what slowed the decline in water levels (Table 2).

	1980-1995 avg yearly precipitation in inches	1980-1995 departure from 30-year avg in inches	1996 precipitation in inches	1996 departure from 30-year avg in inches
Colorado	17.5	+1.3	19.0	+2.0
Kansas	22.6	+1.3	26.7	+5.4
Nebraska	23.3	+1.5	24.1	+2.3
New Mexico	17.6	+1.3	15.4	-1.0
Oklahoma	21.3	+1.3	25.3	+5.3
South Dakota	20.3	+2.0	22.4	+4.0
Texas	20.1	+1.2	18.6	-0.3
Wyoming	15.6	+1.1	13.6	-0.9
<b>High Plains Aquifer</b>	<b>21.2</b>	<b>+1.4</b>	<b>22.0</b>	<b>+2.2</b>

Table 2: Precipitation and Changes in Precipitation in the High Plains Aquifer, 1980-1997



3. Read “*Plains Farmers Learn from Past as Aquifer Depletes*” (attached). The two farmers quoted in the article have different points of view on using the water from the High Plains (Ogallala) Aquifer. Summarize their arguments in the space that follows. Which argument do you think is fairer?

4. Complete the following table and answer the question that follows.

State	1950-1980 water level change in feet	1980-1995 water level change in feet	1950-1980 Change PER YEAR in water level in feet	1980-1995 Change PER YEAR in water level in feet
Colorado	-4.2	-5.9		
Kansas	-9.9	-8.0		
Nebraska	0.0	+2.3		
New Mexico	-9.8	-4.0		
Oklahoma	-11.3	-2.7		
South Dakota	0.0	+3.1		
Texas	-33.7	-7.4		
Wyoming	0.0	-1.5		

**Changes in Water Levels in the High Plains Aquifer, 1950-1995**

What do the data that you just calculated (per year changes) give you?

5. What other information or data would you need to explore the issue of irrigation versus above average rainfall further?

## Plains Farmers Learn from Past as Aquifer Depletes

by Debbie Elliott

NPR, *All Things Considered*, August 11, 2007



Kate Davidson, NPR

Dryland farmer Bill Mai sits at the well his father drilled into the Ogallala Aquifer in 1948. Since then, the water table beneath the Mai Farm has dropped about 80 feet.

[All Things Considered](#), August 11, 2007 · About a quarter of a million people left the Great Plains during the Dust Bowl of the 1930s, but for the families who stayed, hope lay in a massive aquifer of clean, fresh water right beneath their feet.

The Ogallala Aquifer, which lies under parts of eight states, was like lifeblood for Great Plains farmers who today pump billions of gallons of water from the underground supply to their crops annually. But widespread use of the Ogallala for irrigation is depleting the aquifer in some areas, leaving farmers with yearly decisions about how to use the water and how to conserve it for future generations.

### **An Enormous Resource**

The Ogallala, a vast formation of sand and gravel, is part of a larger aquifer that once held more than 3 billion acre-feet of water, roughly the equivalent of 3 billion acres of land covered in a foot of water. The aquifer's depth varies by state. In parts of Nebraska, there are still 1,000 feet of water saturating its sand and gravel. But in parts of Texas, Oklahoma and Kansas — states that had less water to begin with — the water table has dropped steeply.

Just outside of Colby, Kan., farmer Lon Frahm uses 38 sprinklers to pump water over 5,000 acres of farmland. His center-pivot irrigation system sprinkles 600 gallons per minute on his corn, making up for where Mother Nature leaves off in a semi-arid zone that gets less than 20 inches of rain a year.

Frahm estimates that he pumps around 2 billion gallons of water from the Ogallala annually. In recent years, his wells have dropped about 6 inches, though in years when there is more rain, he uses less water. As a member of the Kansas water board, Frahm regularly wrestles with the question of what to do as the water table drops.

"We've got some people that use the analogy of oil. What's the difference between taking oil from the ground, and taking water from the ground? It's a resource not doing any good down there," Frahm says, citing a view he often hears from other farmers.

Kansas has created local groundwater management districts to manage water use in the state. But some believe that only a wholesale shift in the economy can reduce the demand on the Ogallala.

### **Improving on the Past**

Some farmers are returning to dryland farming — raising crops without irrigation — rather than depending on the Ogallala for hydration. Bill Mai, a farmer in Sharon Springs, Kan., was born in the middle of the Dust Bowl. During the worst of the Dust Bowl, his parents considered moving away, but they mustered the will to stay, and discovered they could tap the Ogallala for their crops.

But Mai realized years ago that there wouldn't always be enough water for him to irrigate, so he modified pre-Dust Bowl farming methods to cultivate his wheat and corn crop. In the early 20th century, farmers plowed millions of acres of the Great Plains to plant wheat. Their plows freed soil from the roots of native grass that had evolved over millions of years to hold the earth in place. Once the wheat withered away, the loose soil blew black blizzards across the Plains. At its peak, the Dust Bowl covered 100 million acres.

Rather than repeat the mistakes of the past, Mai now protects his soil by leaving stubble from previous harvests intact. Rain falls on the ground cover, creating a spray that drains down into the soil. He says that dryland farming will help conserve the Ogallala for his grandchildren.

"The water here is probably the most valuable natural resource that we have for every purpose that there is," he says. "If you're looking at economic development and we don't have water, you're dead in the water. Literally."

## Appendix 2: Teacher and Principal Interview Protocols

Table 1. Discussion Guide for Teacher Interviews during TWD Field Test

- 
- Interviewer introduction – discuss confidentiality of responses
  - Thinking about the module you just taught ...
    - Give me an overview of how you taught the module (Evaluation Question 1)
      - Describe a “typical” lesson
    - In your own words, what were the module objectives? (Evaluation Questions 1 & 2)
    - What worked and what did not work? (Evaluation Question 1)
    - Were your students engaged? (Evaluation Questions 2 & 3)
      - Examples
      - Is this more or less than usual?
    - Did your students increase their understanding of data literacy and proportional literacy compared to prior to the lesson? (Evaluation Question 2)
      - Examples
      - How are you assessing it?
    - Did your students increase their understanding of the content you were teaching them? (Evaluation Question 3)
      - Examples
      - How are you assessing it?
    - What technology did you use to teach the lesson? (Evaluation Question 4)
      - If used any, what impact did it have?
      - If used any or did not use any, what else would you have wanted to more effectively teach the module?
  - What might have helped you to better teach this module to your students? (Evaluation Questions 1 and 2)
  - Thinking about expanded distribution of the materials so that it can be taught in more schools, what suggestions do you have for preparing other teachers to be able to implement the module? Why? (Evaluation Question 5)
    - Any Supplemental materials/resources needed?
    - Any additional instructions for teachers?
    - Any other specific changes to the modules?
  - Possible probe for their level of preparedness to teach the module – depends on answers to last two questions (Evaluation Questions 1 & 5)
  - Any other comments?
-

Table 2. Discussion Guide for Administrator Interviews during the TWD Field Test

- 
- Interviewer introduction – discuss confidentiality of responses
  - Thinking about the series of modules your four 7<sup>th</sup> teachers taught this year...  
(Specifics to consider might be: a) two week blocks of instruction per subject; b) sequence of modules; c) prepared instructional materials; d) notebook use by students; e) computer lab/computer carts.)
    - What worked and what did not work? (Evaluation Questions 1 & 5)
      - How do you know this?
      - Give examples
    - How did your school's participation in this project affect teachers and/or students behaviors? Did you observe changes? (Evaluation Questions 1, 2, & 3)
      - Examples
      - What do you base these perceptions on?
    - What technology or other resources did you have to make available to your teachers? (Evaluation Question 4)
      - What impact did it have on other teachers in your building?
      - Do you think you could have used other or additional resources?
    - Would you encourage your teachers to utilize these modules again next year? (Evaluation Questions 1, 2, 3, & 5)
      - As a group or individually?
      - Why or why not?
  - How is this program similar or different from what your team teachers already do? (Evaluation Questions 1 & 5)
  - Thinking about expanded distribution of the materials so that it can be taught in more schools, what suggestions do you have for preparing other school administrators to be able to implement these modules? (Evaluation Question 5)
    - Advantages?
    - Challenges?
    - Any supplemental materials/resources needed?
    - Any additional preparation for teachers?
  - Any other comments?
-