Using ‘Rules of Thumb’ Practices to Enhance Conceptual Understanding and Scientific Reasoning in Project-based Inquiry Classrooms

Michael T. Ryan & Janet L. Kolodner
Learning By Design, Georgia Institute of Technology, 801 Atlantic, 276 CRB, Atlanta, GA. 30332
Tel: 404-894-1987, Fax: 404-894-5041
Email: mryan@cc.gatech.edu, jlk@cc.gatech.edu

Abstract: With foundations in scientific argumentation/discourse literature and transfer literature, this study describes the potential of a new ritualized and repeated classroom activity, the *Rules of Thumb practice*, in developing the conceptual understanding, scientific reasoning, and transfer ability of physical science students in project-based inquiry classrooms (e.g., Learning By Design). Teachers employ an experimental Rules of Thumb practice with more or less fidelity to develop student *science talk* (defined as the skill or act of communicating and explaining, both in written and/or verbal form, the science concepts and principles within a context in an abstract, generalized form) using scaffolded, iterative instructional practices. Comparison and experimental classrooms completed two post-treatment writing assessments, which were coded and analyzed. This paper presents the results of that analysis and reports that the Rules of Thumb practice may have an effect in developing conceptual understanding, scientific reasoning, and transfer ability and that teacher implementation of the Rules of Thumb practice does indeed affect student outcomes.

Introduction

Researchers and educators in the science education community have developed curricula, methods, and practices to improve science literacy, focusing on conceptual understanding, scientific reasoning, and transfer. The need for improvement in these areas has been highlighted by a variety of assessments of U.S. science education (e.g., NCES, 2001). One way researchers and teachers have addressed these needs is through the development of project-based inquiry units (NRC, 1996). There have been significant successes reported from these endeavors, but one area that remains difficult is helping learners make connections between the phenomena they are experiencing and the scientific principles that explain them (Roth, 1997). Indeed, this has been a difficulty we have encountered in our Learning by Design (LBD) classrooms (Kolodner et. al., 2003).

But how might one make that happen in a project-based inquiry setting – one where learners are investigating phenomena and their effects in order to achieve some challenge? Jimenez-Aleixandre & Rodriguez (2000) conducted research in an inquiry and problem-based, high-school genetics class where teachers created an environment in which students were encouraged to socially defend their positions with data and science concept principles. This research found that this method created students who: 1) engaged in scientific argument; 2) asked for each other to explain their arguments; and 3) developed warrants, i.e.- reasons based in principles that make the explicit connection between data and a conclusion or claim. While a teacher, the first author of this paper discovered a way to do this in his classes by enhancing a public Rules of Thumb practice that focused on explanation and iteration of explanations. This seemed to be a success, so we decided to investigate its effects in a more formal way.

LBD’s founding principles and learning models (problem-based learning, project-based inquiry, and case-based reasoning) contribute to the founding of this study (Kolodner, et. al., 2003). Evaluations of LBD show that students improve their conceptual understanding of scientific concepts, their scientific reasoning, and collaboration behaviors. The best practices, however, for helping students develop scientific reasoning, content understanding, and promoting transfer were not quite clear. This paper reports on an investigation of the effects of enacting one LBD classroom practice (*the Rules of Thumb practice (RT)*) in three different ways – LBD’s original way, a new way that focused upon the development of science talk (defined by this author as the skill or act of communicating and explaining science concepts and principles governing a context in an abstract, generalized form), and something in between. The new RT practice, when implemented with high fidelity, promotes greater use of science content to explain experimental results and justify project decisions. It focuses on moving students from discussions about the specifics of their project experiences to more abstract and generalized articulation of science concepts that explain them. Data shows that this results in greater proclivity by students to recognize the appropriateness of engaging in science talk, to do so with more depth, and to use science content knowledge and scientific reasoning skills in a new situation. Our research was done in three Atlanta-area middle schools.

The abstraction of science concepts, ideas, principles, and norms of practice is an important goal of science education (NRC, 1996). The transfer literature, however, tells us this is difficult to achieve because the appropriate recognition and use of knowledge outside of the context in which it is learned is rare (Bransford, et.al., 1999). The new RT practice was engineered with this in mind. The new protocol might serve as a model for classroom practice that develops student abstraction of science concepts and scientific reasoning skill.

Background - LBD's "Original" Rules of Thumb

Learning by Design (LBD) is a project-based inquiry curriculum approach designed to improve performance in science literacy at the middle school level for all students. In LBD curriculum units, student groups design and develop a *design artifact* (an actual object or model) that achieves the design challenge. In *Vehicles in Motion* (*VIM*), a unit covering forces and motion, students build, experiment with, and redesign a model car and its propulsion mechanisms. Students either design or re-engineer the artifact to improve its performance, allowing them to meet a challenge put forth by the curriculum unit and their teacher. LBD students, in groups, experiment with features of the design artifact to learn about their effects on the performance. The information gained from these experiments informs and drives design decisions and future versions of the artifact. Experimenting with the artifact's features or aspects (i.e., scientific variables) provides students the opportunity to experience first-hand the science concepts governing the performance or purpose of the design artifact. Furthermore, this opens the door for teachers to connect the science concepts to design decisions learners must make to meet the challenge.

Central to LBD's methodology is the use of repeated public *practices* (Kolodner, Gray, & Fasse, 2002) where teachers and students engage in discourse to scaffold development of science content understanding, science inquiry skills, collaboration skills, and scientific reasoning (Kolodner, et al., 2003). The RT practice follows a round of experiments conducted by the student groups formed in the class. Following experiments, a Rules of Thumb session is held in which each student group presents a recommendation for future designs (i.e., their Rule of Thumb) based on their experimental data. Those chosen by the class as justified by experimental results are recorded on a poster. During later RT sessions, after students have attempted to apply those rules of thumb, the class returns to the RT poster and discusses the merits and meanings of all the current and prior RTs suggested. Students need the results of all experiments done by groups in the class, and hence an understanding of the full set of RTs to succeed at their challenge.

This practice and the public record it produces served as vehicles for converging on and recording useful design information and to remind students of what they should take into account in further designing. But early implementations produced RTs that were not scientifically rigorous. Often, students were making good observations and recommendations, but they didn't go deeper than a surface design suggestion or beyond fabrication techniques (e.g., "use tape to secure bearings to the car because it can be removed and reapplied more easily than glue") or feature settings (e.g., "keep the wheels positioned far away from the chassis so that they do not rub on the chassis"). Their RTs were never decontextualized, and students rarely explained the effects of the variable using science content. To promote and scaffold this skill, the first author (LBD teacher, 1999-2000) formulated an RT template:

> **"When** <u>(describe the action, design, or choice you are working within)</u>, **use/connect/build/employ/measure** <u>(list your suggestion or method)</u> **because** <u>(list or supply the science principle or concept here that backs up your suggestion)</u>**."**

The introduction of this template produced substantial changes to the types of RTs observed. RTs changed from statements such as: *"Use straw bearings because they allow the car to travel farther than any of the other bearings"* to *"When choosing a bearing for the coaster car, use the straw bearings because the straw has very low friction and will not cause the car to slow down as much as the other bearings because other bearings change the net force a lot"*.

The “New” Rules of Thumb Protocol

The ubiquitous theme of the new RT practice is overt scaffolding of opportunities for students to (1) develop their science talk faculty and (2) argue and justify decisions, publicly, with science concepts learned during VIM. The goal of the practice is to have students, in a RT session, create RTs that focus not only on the design idea or solution itself, but also on the science governing the success of the idea or solution. In addition, the RTs, including their scientific explanations, are iteratively revisited and revised as the class moves through the unit. The practice has several important characteristics: (i) it expects and fosters student engagement in science talk and science argumentation/reasoning; (ii) it creates opportunities for students to engage in science talk and science argumentation/reasoning, both in written and verbal formats; (iii) it targets student-centered, versus teacher-centered, science talk and science argumentation/reasoning; (iv) it emphasizes use of scaffolding tools to develop science talk and science argumentation/reasoning, and (v) it encourages and rewards peer critiquing and Socratic discussion of students’ science talk and science reasoning.

The predicted outcome of the new RT practice was that students would demonstrate greater proficiency in science talk, scientific reasoning, and transfer. Specifically, they would (i) use science concepts and principles learned during the unit to explain, justify, and reason decisions, problems, and solutions at a higher rate; (ii) articulate these science concepts and principles in a more abstract, generalized voice or manner; demonstrating greater conceptual understanding and greater tendency to reason scientifically; and (iii) demonstrate higher quality transfer of content knowledge and reasoning skills and at a higher rate.

To articulate the requisite scaffolding (Holbrook & Kolodner, 2002), the LBD team designed a special worksheet, entitled *My Rules of Thumb* (Figure 1) to help students record and develop RTs according to the identified protocol. Teachers help the students create a public RTs list from the rules that different groups suggest.

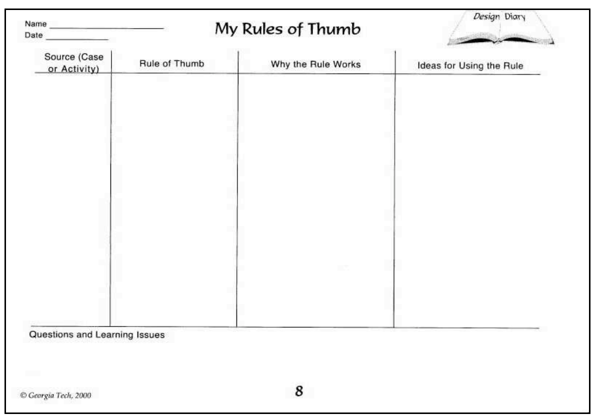

Figure 1 – *My Rules of Thumb* Design Diary sheet.

The new practice dictates iterative refinement to edit or eliminate poor RTs. The *My Rules of Thumb* sheet is a *dynamic* record of rules that develop over time with experience, and it is revisited during other LBD activities. This allows students to practice with and engage in *science talk* and discourse more often. Table 1 emphasizes the importance of RTs by highlighting (in italics) when they are used in LBD’s *Vehicles in Motion* curriculum sequence.

Method

Research design

The nature of the curriculum, the time constraints placed on teachers by their school districts, and the similarities between the cohorts (Kolodner, et.al., 2003) dictated a quasi-experimental, static group comparison design (Salkind, 1997). Four 8th-grade teachers who were implementing VIM in metro Atlanta suburban schools participated in this study. The cohorts involved were of a similar demographic, and each teacher implemented the unit at approximately the same time during the 2000-2001 academic year. The two teachers in the experimental group (T1and T2) were to implement the new RT protocol, and two comparison teachers (T3 and T4) were to implement the old one. The teachers trained in their respective protocols during a summer workshop and prior to the unit’s implementation. However, T4 fell ill and dropped out of the implementation, and the two experimental teachers each implemented the new practice protocol differently. Thus, we were left with three groups to compare.

Table 1: Annotated LBD/VIM Annotated Sequence with New Rules of Thumb Practice Elements

1. Challenge Introduction, Build, & Mess About: Problems specifications, constraints, and criteria identified. Build basic version of the artifact. Examine and Mess-About with the basic design; students share observations.
2.Whiteboard the Challenge: Class reviews observations, constructs a Whiteboard to organize ideas and variables to test.
3. Design Experiments: Students identify specific variables to test, plan experiments.
4. Conduct Experiments: Students create and conduct experiments to test variables identified.
5. Poster Session: Students conduct a Poster Session to present their experiment(s) to the class.
6. Rules of Thumb Presentation: Presenting group generates a RT regarding their variable to help students plan future designs. RTs are recorded and discussed. Students record RTs on the My Rules of Thumb sheet, knowing that it may be edited later. Teacher models science talk.
7. Science Concept Instruction: The class's attention focuses on the science issue at hand. The teacher emphasizes and models science talk: Rephrasing students answers and asking students to rephrase answers using the science vocabulary. Teacher emphasizes heavily the relationship between the abstract science principles and the examples students have generated, referring to the challenge or problem often in explanations.
8. Conduct More Experiments, Poster Sessions, Rules of Thumb Session (Cycling steps 3-7): Before starting the next round of experiments, students review the RTs to identify and rewrite any misstatements or misconceptions based on their recent research. Students once again discuss and deliberate over RTs as a class. This time the class is more prepared to fill in the Why the Rule Works column because of their exploration of the science concepts. Teachers help students reword justification or explanation (rephrasing as a question) to model the use of abstract decontextualized language. The teacher models science talk both in discussion and when writing RTs. As steps 3-7 are repeated, a Rules of Thumb Session occurs, and the My Rules of Thumb sheet is edited every time.
9. Plan Final Design, Pin-Up Session: When the class completes the experimenting phase, student groups review RTs, review experimental results, and consult with other groups to plan final design of artifact. Each group creates a Pin-Up: a sort of blueprint of their final design idea. The class convenes to share Pin-Ups.
10. Build & Test, Gallery Walks: Groups build and test their final design. Data is collected and Gallery Walks occur. During redesign, students may verify RTs or sometimes, they may discover flaws in their RTs. Once again, the teacher uses guided questioning and discussion (perhaps even more demos, science concept research, homework assignments) in an attempt to have students develop more valid RTs and better articulate science concepts within a context as evidenced by their RTs. This step highlights the priority and encouragement of an iterative development of science. Again, the teacher models science talk both in discussion and when writing RTs.
11. Final Presentation: Students present their final artifact and attempt the challenge(s).
12. Review Challenge, Whiteboard: The class reviews the limitation of the artifact type, in general, and identifies what is further required to complete the challenges more successfully which leads into the next module. Teacher and students review the Rules of Thumb, identifying those that seemed to have the most influence on successful designs, i.e. – those that met the challenge.

Instruments

The post-tests were assignments that were part of the curriculum unit. In the first assignment (A1), each student wrote a *Product History*, a detailed explanation of the evolution of his/her model vehicle. These papers discuss problems encountered, design changes, experiments on design features, and explanations of design choices, both successful and unsuccessful. As students describe the features, problems, and aspects of their vehicles, there is a tremendous opportunity for students to relate the science they have learned, especially since the assignment explicitly asks for them to do so. In the second post-test assignment (A2), students create and write recommendations for a team of scientists who are planning to design and build a vehicle to explore Antarctica. The assignment details the conditions the vehicle must endure, and students are asked to comment on several areas of the vehicle in their recommendation. Again, students can identify the concepts at work, explaining how and why those concepts govern the feature discussed.

Quasi-experimental, static group comparison design (Salkind, 1997) lacks a pretest. A pretest was not possible for this study. It was very difficult to create pretest assessments similar to the end of unit assessments that could target the existing knowledge and abilities measured in this study. Additionally, because of constraints on teacher-instruction time and other aspects of implementing LBD that year, the teachers were unwilling to participate in pretest activities or have students do work that was in excess of that already in the unit. Finally, it was the teachers' wish that students complete only written assignments already existing in the LBD unit (i.e. – Product History and Antarctica Car Recommendation) for our study. All of these factors, in aggregate, eliminated the option of a pretest.

Predictions

The written assignments offered students the opportunity to explain and justify design problems, features, solutions, experiments, or ideas based upon their project experience. A1 offered a view of each student's ability to engage in science talk, demonstrating propensity and prowess in their conceptual understanding, scientific reasoning, and ability to apply what they had learned within the context it was learned in. Use of science content and science talk skills with this assignment could be considered evidence of near-transfer. A2 offered more of a far-transfer opportunity to apply science content knowledge to a new, unfamiliar situation while still offering students the opportunity to explain and justify with science talk.

Results

Teacher Practices

If we place the variations of the RT practice on a continuum of treatment fidelity, T1's implementation can be defined as 'high-fidelity'. T1 not only adhered to the prescribed treatment, but she often enhanced and exceeded the requirements of the experimental treatment. T2 deviated from the prescribed experimental treatment, but not so much that her implementation was void of elements of the new RT practice. T2 evolved into a quasi-experimental teacher, a comparison teacher to both T1 and T3. T2's fidelity might best be characterized, in terms of this continuum, as low-fidelity. T3, the comparison teacher, implemented the old RT practice where students develop and revisit RTs together but do not use them to connect science concepts and design challenge activities explicitly or to develop science talk fluency.

> T1 (high-fidelity new-protocol group) frequently modeled science talk, engaged students in speaking and writing science talk, and had students edit their science talk during peer-feedback situations. T1 also expected students to engage in science talk on a regular basis during other parts of the project-based sequence of activities.
>
> T2 (lower-fidelity new-protocol group) provided students fewer opportunities than prescribed to develop science talk skills. When rules of thumb were discussed, the moments were usually teacher-centered in nature. Students had discussions about the science, trying to link their RTs to the science; however, they never recorded these justifications and explanations in a deep manner on their *My Rules of Thumb* sheets. Furthermore, T2 employed the strategies and methods during only some of the points in the sequence as instructed during training.
>
> T3 (old-protocol group) implemented the comparison treatment as anticipated. Even when students afforded her the opportunity, T3 did not provide the scaffolding and iterative development of the connections that the research literature suggests is necessary. Rather, T3 recorded the suggested RTs and then allowed students to decide, individually, whether each was important enough to consider when designing their artifacts.

A full commentary and review, including classroom transcripts, of the practices and variations of the teachers are available in Ryan (2003).

Coded Student Assignments

Because the data are non-parametric, chi-square analyses, where expected frequencies are compared to observed frequencies, was most appropriate (Shavelson, 1981). Null-hypotheses were that there would be no predictable relationship between teacher treatment and category classification distributions. A new coding scheme assessed the written assignments to determine student use of science principles in explanation. The student assignments were coded and reliability was established.

Each time a student discussed a feature or problem in the assignment, the feature or problem was labeled as an *Event*. Each Event in each assignment was coded across three critical categories: Justification, Form, and Depth. The results for each coding category are shown in Tables 2 – 10, which provide the rates of occurrence for each teacher in each coding category and the chi-square analyses of the outcomes.

Justification

This category classifies Events according to whether or not the student chose to explicitly justify or explain the Event with a science concept.

Table 2 – Justification Results for Events Coded in A1, Percentage of Occurrence

	T1-High Fidelity	T2- Low Fidelity	T3-Old Practice
Not justified via science concept	20%	54%	76%
Justified via science concept	80%	46%	24%

The chi-square analysis does support the rejection of the null-hypothesis for: T1 v. T3, with $\chi^2(1, N = 1332) = 300.1$, $p<0.001$, T2 v. T3, with $\chi^2(1, N = 1332) = 29.4$, $p<0.001$, T1 v. T2, with $\chi^2(1, N = 1332) = 97.4$, $p<0.001$

Table 3 – Justification Results for Events Coded in A2, Percentage of Occurrence

	T1-High Fidelity	T2- Low Fidelity	T3-Old Practice
Not justified via science concept	4%	7%	50%
Justified via science concept	96%	93%	50%

The chi-square analysis does support the rejection of the null-hypothesis for: T1 v. T3, with $\chi^2(1, N = 233) = 52.1.1$, $p<0.001$, T2 v. T3, with $\chi^2(1, N = 233) = 21.1$, $p<0.001$

Table 4 –Examples: T1 and T2 students tended to provide science-based Events more often than T3 students.

T1	"The car had straws for bearings and CDs for wheels. These were chosen because they reduce friction with the floor as the car is rolled."
T2	"The wheels we chose were CDs because they were light and created a low amount of friction on the floor."
T3	"The wheels we chose were CDs. They were large and worked well on our coaster car."

Form

Each Event that was categorized as justified by a science concept was coded for Form. Form codes whether the Event cites only the science term (e.g., friction) or if it also discusses the principle behind the term (e.g., a force created on the surfaces of two objects in contact).

Table 5 – Form Results for Events Coded in A1, Percentage of Occurrence

	T1-High Fidelity	T2- Low Fidelity	T3-Old Practice
Citing only science term, no principle	61%	80%	85%
Citing science principle	39%	20%	15%

The chi-square analysis does support the rejection of the null-hypothesis for: T1 v. T3, with $\chi^2(1, N = 769) = 13.7$, $p<0.001$, T1 v. T2, with $\chi^2(1, N = 769) = 9.4$, $p<0.005$

Table 6 – Form Results for Events Coded in A2, Percentage of Occurrence

	T1-High Fidelity	T2- Low Fidelity	T3-Old Practice
Citing only science term, no principle	61%	80%	85%
Citing science principle	40%	22%	17%

The chi-square analysis does support the rejection of the null-hypothesis for: T1 v. T3, with $\chi^2(1, N = 172) = 6.5$, $p<0.025$, T1 v. T2, with $\chi^2(1, N = 172) = 3.9$, $p<0.05$

Table 7 – Examples: T1 students often cited science concepts explaining the principle behind it.

T1	"We started off using one [balloon] engine with a single balloon attached. Then we started to attach more engines because, as Newton stated in his second law, the more net force on the car, the more acceleration it will have because the extra engines do not weigh that much at all. More engines gave us more overall force and speed."
T2	"We created [with more balloon engines] more net force and made the car go faster."
T3	"We added a third engine because of Newton's 2nd law."

Depth

Each Event that was coded as having been justified via science concept was coded for Depth. Depth signals whether the science concept in the Event is portrayed in an abstract, generalized format or in terms of the context of the feature or problem.

Table 8 – Form Results for Events Coded in A1, Percentage of Occurrence

	T1-High Fidelity	T2- Low Fidelity	T3-Old Practice
Contextual reference of science	82%	98%	95%
Abstract reference of science	18%	2%	5%

The chi-square analysis does support the rejection of the null-hypothesis for: T1 v. T3, with $\chi^2(2, N = 769) = 34.3$, $p<0.001$, T1 v. T2, with $\chi^2(2, N = 769) = 18.4$, $p<0.001$

Table 9 – Form Results for Events Coded in A2, Percentage of Occurrence

	T1-High Fidelity	T2- Low Fidelity	T3-Old Practice
Contextual reference of science	89%	100%	100%
Abstract reference of science	11%	0%	0%

The chi-square analysis cannot be completed because the expected frequency for any comparisons < 5, with df ≥ 2.

Table 10 –Examples: T1 and T2 students tended to provide science-based Events more often than T3 students.

T1	"In our balloon car we also tripled and double layered the balloons. We did this because the air coming out would have more force on it because there was greater tension. Therefore, according to Newton's 3rd law, each of these object experience equal and opposite forces. The push on all the air molecules would be the same size force on the balloon in the opposite direction the air goes."
T2	"We double layered our balloon engines because to create a bigger force back on the balloon (3rd Law) and make the car go farther."
T3	"We layered our balloons and this provided more force for our car"

Discussion

As the National Science Education Standards indicate, inquiry methodology and curriculum promote quality science education (NRC, 1996). There are, however, many aspects to inquiry-based learning. One aspect is the engagement of students in scientific argumentation and scientific discourse. Research on involving students in this type of behavior suggests positive effects (e.g. – Driver, et. al., 2000). Argumentation, as a practice, is essential in science communities as a way to assess and evaluate the work of scientists. Participation in student-centered constructive discourse and argumentation about science findings and principles improves conceptual understanding and encourages reasoning from science.

The new RT practice targets *science talk* –communicating and explaining science concepts and principles governing a context in an abstract, generalized form. Students often cannot use the abstract, generalized language of the science domains easily to discuss, explain, justify, or reason real and hypothetical events (Roth, 1997). LBD units includes a variety of public activity structures designed to help teachers and students engage in science talk (Kolodner, Gray & Fasse, 2002). Closely related to the idea of student engagement in scientific argumentation is the lack of student experience in using science talk to communicate ideas. Students need guidance, scaffolding, practice and reflection to develop this skill (Roth, 2001; Driver, et. al., 2000). Furthermore, written expression of science conceptual understanding posses even more challenges than verbal expressions (Bereiter & Scardamalia, 1987). This research informed the design of our more scientifically-enhanced RT practice. Adding this ritualized way of engaging in such science talk, we thought, would promote more and higher-quality science talk in LBD classrooms.
Students learning in the high-fidelity enactment engaged in scientific discourse and argumentation (1) often, (2) as an explicit act to develop understanding and socially construct knowledge, and (3) to practice reasoning as scientists would. Previous research predicted that classrooms focusing on these concepts would be impacted in the way T1's and, in a more limited way, T2's classes were (Jimenez-Aleixandre & Rodriguez, 2000). Experimental students confirmed these studies, by demonstrating firmer conceptual understanding and a tendency to reason scientifically on A1. High-fidelity students are offered more opportunities and practice for speaking and writing science talk. As a result, these students are comfortable expressing science in a sophisticated way, as Roth (1997, 2001) and Chi, et. al. (1991) suggest. The results might indicate that scaffolded, iterative writing of RTs promotes higher science talk faculty, which, perhaps, impacts conceptual understanding and reasoning.

The results also may suggest impacts on transfer. In terms of scientific reasoning, T1 and T2 students scored similarly on A2, the more-far transfer assignment. These students chose to justify design decisions citing science concepts at very high rates, 96% (T1) and 93% (T2), compared to T3's 50% rate. These students seem to recognize the importance of doing so, and perhaps they look to identify science concepts in new contexts more often and more easily than do the comparison students – a sign of increased *situational transfer* (Bereiter, 1995). These T1 and T2 results demonstrate a propensity to reason with science in a way that parallels their actions and behaviors for RT development during VIM. Results also suggest T1's students tend to develop explanations and justifications that not only identify relevant concepts, but also include deeper, abstract explanations of it's the connections between those concepts and features of the designed artifact. The new practice focused on this skill, and T1-students recognize and employ it outside of the RT contexts where they were expected to use it (i.e. – during RT sessions or experiment presentations). It is possible that T1's focus on science talk development during times other than the RT sessions influenced student outcomes, and study isolating science talk occasions may be pertinent. Research on transfer also discusses that if and when transfer occurs, it requires a long time and many experiences (Bransford, et. al., 1999). Learners need multiple experiences to start making more general connections. Here, all of the LBD students spent 10-12 weeks learning about *force* and *motion* and investigating aspects of these concepts from multiple angles. Aside from the RT practice, they engaged in the same sequences of activities. Our results show that the number of experiences and time spent engaging with a concept, by themselves, are no guarantee of development of deep understanding. The students who engaged in the experimental RT practice methodically developed their understandings and expressions of science concepts, formulating their own explanations and representations. Their more consistent ability to use those concepts in new situations was a result of multiple public and private interactions with the concepts and the iterative development of generalized understanding of the concepts.

It is important in this final discussion to acknowledge the amount of scaffolding, training, and tools provided to the experimental teachers. Newton, et al., (1999) and Geddis (1991), suggest that teachers are ill-prepared to handle student/peer centered discourse and argument, and they need guidance and time to develop these skills. Our experience and research, here, supports these claims. Teachers, themselves, need structure, coaching, and practice to develop student understanding and reasoning via science talk and discourse. Perhaps creating a more generalized but consistently used RT methodology for teaching science to would well serve project-based inquiry teaching and research. The methods presented here are a step in that direction, but more input and development is needed. Ultimately, the motivation for giving abstract, general responses rooted in science is difficult to assess without more comprehensive measures. Perhaps, there could be another Rules of Thumb study where students have other opportunities, beyond the written assignments, to reveal their understanding and motivations. The use of student interviews and performance assessment activities might offer researchers the ability to make more confident claims regarding Rules of Thumb practice.

References

Bereiter, C (1995). A Dispositional View of Transfer. In McKeough, A., Lupart, J.L., Martini, A., *Teaching for Transfer: Fostering Generalization in Learning*. Mahwah, N.J.: Lawrence Erlbaum Associates, pg. 21-34.

Bereiter, C. & Scardamalia, M. (1987). An attainable version of high literacy: Approaches to teaching higher-order skills in reading and writing. *Curriculum Inquiry*, 17 (1), 9-30.

Branford, J. D., Brown, A. L., Cocking, R. R., eds. (1999), *How People Learn: Brain, Mind, Experience, and School*. Washington D.C.: National Academy Press.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P. & Glaser, R. (1989). Self-Explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 13, 145-182.

Driver, R., Newton, P., & Osborne, J. (2000). Establishing norms of scientific argumentation in classrooms. *Journal of Science Education*, 84, 287-312.

Geddis, A. N. (1991). Improving the quality of science classroom discourse on controversial issues, *Science Education*, 75 (2), 169-183.

Holbrook, J. & Kolodner, J.L. (2000). Scaffolding the Development of an Inquiry-Based (Science) Classroom. In Proceedings of the International Conference of the Learning Sciences 2000 (ICLS) , pp.221-227. Last retrieved February 17, 2004 from Web site: http://www.cc.gatech.edu/projects/lbd/htmlpubs/scaffolddevelop.html

Jimenez-Aleixandre, M. P. & Rodriguez, A. B. (2000). "Doing the Lesson" or "Doing the Science": Argument in high school genetics. *Science Education*, 84 (6), 757-792.

Kolodner, J.L., Crismond, D., Fasse, B., Gray, J., Holbrook, J., Puntembakar, S. (2003). Putting a Student-Centered Learning by Design™ Curriculum into Practice: Lessons learned. *Journal of the Learning Sciences,* Vol.12 No.4.

Kolodner, J.L., Gray, J. and Fasse, B. (2002).Promoting Transfer through Case-Based Reasoning: Rituals and Practices in Learning by Design™ Classrooms. *Cognitive Science Quarterly*, Vol.

National Center for Educational Statistics, NCES (2001) *National Average Science Scale Score Results, Grades 4, 8, and 12: 1996 and 2000.* Last retrieved April, 10, 2003 from Web site: http://nces.ed.gov/nationsreportcard/science/results/natscalescore.asp

National Research Council (1996). *National Science Education Standards*. Washington D.C.: National Academy Press. Choose one or at most two of the National … no room for more here

Newton, P., Driver, R., Osborne, J. (1999). The place of argumentation in the pedagogy of school science, *International Journal of Science Education*, 21 (5), 553-576. choose fewer of the argumentation references; put e.g., before what you choose in the text (e.g., Driver, ….)

Roth, W–M. (1997). Why may students fail to learn from demonstrations? A social practice perspective on learning physics. *Journal of Research in Science Teaching*, 34 (5), 509-533.

Roth, W–M.. (2001). Students' talk about rotational motion within and across contexts, and implications for future learning. *International Journal of Science Education*, 23 (2), 151-179. can you choose one or the other?

Ryan, M. T. (2003) *Using a new design rules practice and science talk development to enhance conceptual understanding, scientific reasoning, and transfer in Learning By Design™ classrooms.* Unpublished master's thesis, University of Kansas, Lawrence. Available at Web site:
http://www.cc.gatech.edu/projects/lbd/pubdate.html#2003

Salkind, Neil J. (1997). *Exploring Researcher*, 3rd edition. Upper Saddle River, N.J.: Prentice Hall.

Shavelson, R. J. (1981). Statistical reasoning for the behavioral sciences. Boston: Allyn and Bacon.

Acknowledgements

Special thanks to Dr. Elizabeth Charles at Learning by Design™, Dr. Jim Ellis of the University of Kansas, and the participating teachers for their contributions to this study and report.