THE USING EVIDENCE FRAMEWORK:
A MODEL OF SCIENTIFIC REASONING

Nathaniel J.S. Brown
Indiana University

Michael Timms
WestEd

Sam O. Nagashima
CRESST/University of California, Los Angeles

Mark Wilson
University of California, Berkeley

Paper presented at the Annual Meeting of the American Educational Research Association,
New York, New York, March 24-28, 2008.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. This material is based upon work supported by the National Science Foundation under Grant Number 0439062

Introduction

This paper describes the work of a collaborative research project called “Using Evidence: An Analysis of US and German Science Teaching and Learning” that was funded jointly by the *National Science Foundation* (NSF Award #0439062) in the United States and by the *Deutsche Forschungsgemeinschaft* (German Research Foundation - DFG) in Germany. The purpose of the project was to build upon existing research that the project team members had conducted on the science curriculum topic of floating and sinking in order to develop an analytic framework and some associated research tools designed to examine how science students and teachers use evidence in the classroom.

This paper, the first in a set of four presented at the 2008 conference of the American Education Research Association, describes the Using Evidence (UE) Framework, a new analytic tool designed to help examine the development of scientific reasoning in elementary, middle, and high school. The analytic framework is also intended to form a basis for the development of instrumentation that will allow other researchers and educators to identify and code teachers’ and students’ use of evidence in teaching and learning. The remaining three papers in the set will illustrate how the UE Framework has been applied to the measurement of different facets of reasoning in science classrooms, specifically with regard to the development of written assessments and the analysis of video recordings of whole-class discussions from classrooms in the US and Germany. By contributing new tools and a rich dataset, the papers in this session help the international science education community further understand how to support student reasoning in reform-based science lessons and, ultimately, to improve students’ understanding of science and the natural world.

The Need for a New Framework of Using Evidence

Students' ability to reason from evidence is considered a major objective of science education reform (American Association for the Advancement of Science, 1993; National Research Council, 1996). In order to participate in arguments about scientific ideas, students must learn how to evaluate and use evidence. That is, apart from what they may already know about the substance of an assertion, students who are scientifically literate should be able to make judgments based on the evidence supporting or refuting a particular claim. The use or misuse of supporting evidence, the language used, and the logic of the argument presented are important considerations in judging how seriously to take a claim or proposition (AAAS, 1993, p. 298). Therefore, in order to be well-balanced and intelligent consumers of scientific hypotheses, students must understand that there should be ample evidence present to determine whether or not hypotheses are valid. This critical thinking skill is crucial for students as they develop more complete understandings of the natural world around them.

Evidence-based arguments form the foundation of scientific thinking, constituting the mechanism by which scientific knowledge is used, tested, and revised (Hempel, 1966; Kuhn, 1993; Schwab, 1962). Reforms in science education that focus on engaging students in the thinking processes and activities of scientists correspondingly include engaging students in evidence-based reasoning (Duschl, 2003; NRC, 1996; 2001). More specifically, reforms emphasize that students should be able to use evidence to "develop and evaluate explanations that help them address scientifically oriented questions, and formulate explanations from evidence" (NRC, 2001, p. 29).

While accumulated research from developmental and cognitive psychology indicates that students’ ability to reason scientifically and to develop conceptual insights has been underestimated, students at all levels still have difficulties differentiating theory and evidence (NRC, 2007). Furthermore, studies have indicated that students rarely base their arguments on evidence (Bell & Linn, 2000). Although many researchers have developed frameworks that analyze the ways in which students use evidence in writing (e.g. Kelly & Bazerman, 2003; McNeill, Lizotte, Krajcik, & Marx, 2006; Sandoval, 2003; Sandoval & Millwood, 2005) or in discussions (e.g., Driver, Newton, & Osborne, 2000; Jimenez-Alexandre, Rodriguez, & Duschl, 2000; Osborne, Erduran, & Simon, 2004; Tytler & Peterson, 2005), there is to date no well-defined framework that makes possible analysis of how students use evidence and how teachers support students in doing so in writing *and* in whole-class talk. Consequently, there are no reliable measurement tools to measure use of evidence.

In order to implement meaningful and valid assessments that will assist teachers to modify instruction in order to help students learn, we must understand how students use evidence to justify their explanations about the natural world. To this end, it is necessary to create a science-specific framework of understanding on the optimal use of evidence by students at different stages of their physical and intellectual development. From this, we can measure teachers’ and students’ actual use of evidence in science teaching and learning.

The Research Collaboration

The research collaboration that developed the UE framework was based in the US and in Germany. In the US, the Using Evidence project involved researchers and graduate students from the Center for Assessment and Evaluation of Student Learning (CAESL) that includes

WestEd, The University of California at Berkeley, The University of California at Los Angeles/CRESST and Stanford University. CAESL's dual mission of improving science learning by focusing on effective assessment, and expanding the research community's knowledge of teachers' and students' use of evidence in science learning, is directly related to this goal. In Germany, the project involved researchers and graduate students from the Max Planck Institute for Human Development, the University of Muenster and University of Munich. Some of the graduate students from the project have subsequently gone on to teaching positions at Indiana University, the University of Colorado, Boulder, and the University of Frankfurt.

The Using Evidence (UE) Framework

Until recently, studies of argumentation in science classrooms have (almost exclusively) relied upon Toulmin's Argument Pattern (TAP) as an analytic tool (Toulmin, 1958/2003). Toulmin describes how an argument consists of a piece of data and a claim, linked together to make a warrant. Evidence then supports this argument in the form of backing (Toulmin, 1958/2003). The relationship among these elements of an argument is represented in Figure 1.

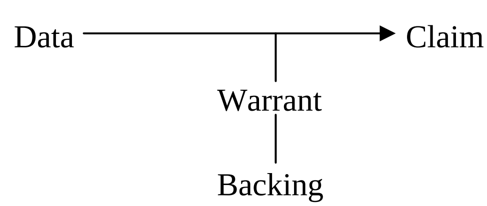


Figure 1. Toulmin's (1958/2003) framework for arguments

To determine when and how evidence-based arguments are taking place, several studies have turned to Toulmin's (1958/2003) foundational framework for the analysis of arguments (e.g. Driver et al., 2000; Jimenez-Alexandre et al., 2000; McNeill et al., 2006; Simon, Erduran, & Osborne, 2006). Despite its frequent use, TAP presents a number of limitations. For instance, difficulties in reliably identifying the parts of an argument that fit into the components of Toulmin's model have forced many researchers to focus on limited aspects of total argument structure (Erduran et al., 2004).

Given these challenges, our Using Evidence Framework draws upon additional theories to further our understanding of argumentation and to inform our conceptualization of student reasoning in science classrooms. Toulmin's specifications for features of an argument (i.e., claims, data, warrants, and backing) provide the backbone for our framework. Additionally, we draw upon Duschl (draft) and McNeill, Lizotte, Krajcik, & Marx (2006) for insight into the role of evidence and argumentation specific to the context of science education. Figure 2 shows the Using Evidence Framework and Figure 3 provides an explanation of all its components.

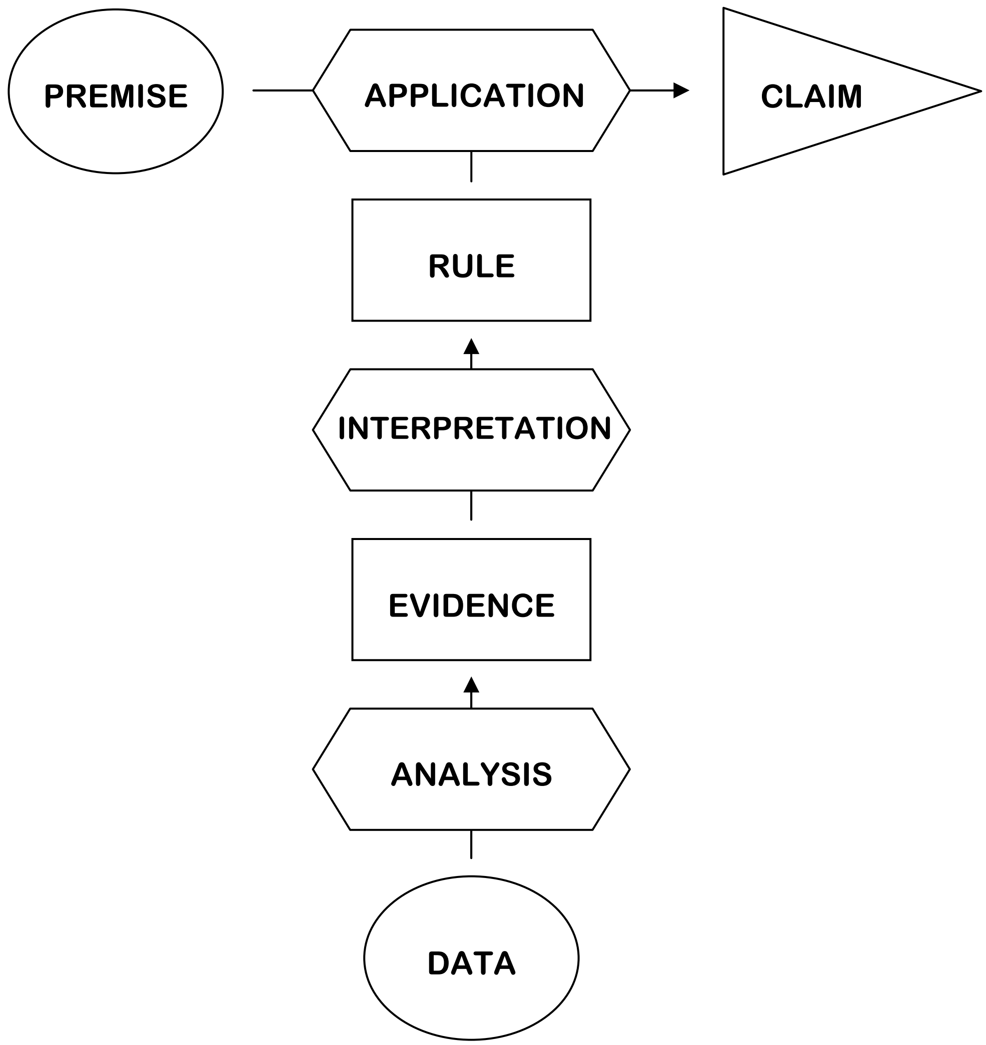


Figure 2. The Using Evidence Framework

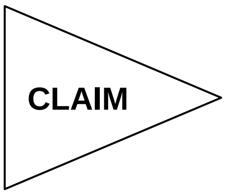
This conceptual model for scientific reasoning distinguishes between two distinct classes of information: Component and Process. A Component, in general, refers to statements that frame and place the argument in context. The model consists of five components, based on Toulmin's (1958/2003) model and Duschl's (2003) framework for assessment of inquiry: Premise, Claim, Rule, Evidence, and Data. *Premise* is a statement describing the specific circumstances acting as an input that will result in the outcome described by the claim. The premise often identifies a specific object and a relevant property (e.g., "this box is heavy"). A

Claim is a statement about a specific outcome or state: what something will do in the future (prediction, e.g., “this box will sink”), has done in the past (observation, e.g., “this box sank”), or is (conclusion, e.g., “this substance is an acid”). It is specific to a single set of circumstances, generally a particular object in a particular time and place. A *Rule* is a statement describing a general relationship between two properties (e.g., “something that is an acid tastes sour”) or a property and a consequence of that property (e.g., “something that is heavy will sink”). This relationship is general in the sense that it is expected to hold even in contexts and circumstances not previously observed. *Evidence* is composed of statements describing a contextualized relationship between two properties or a property and a consequence of that property. This relationship is contextualized in the sense that it describes (or assumes) a specific set of circumstances in which the relationship has been actually observed to be true. *Data* are statements describing a specific relationship between two properties or a property and a consequence of that property. This relationship is specific in the sense that it describes (or assumes) a specific, single event that has been either observed or made up.

The Process pieces of the conceptual model are composed of three parts - Application, Interpretation, and Analysis. *Application* refers to the process by which the rule is brought to bear in the specific circumstance(s) described by the premise. It establishes the probability or necessity of the claim, often by logical deduction in simple cases. *Interpretation* is the process by which multiple pieces of evidence are compared and integrated or synthesized. It establishes the probability or necessity of the rule, often by generalization in simple cases with only one piece of evidence. *Analysis* is the process by which multiple data are compared and integrated or synthesized. It establishes the probability or necessity of the evidence, often by extrapolation in simple cases.

Description of Parts

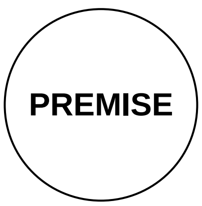
Output



A **CLAIM** is a statement that describes what something either (a) *will do* in the future (called a **prediction**, e.g., “this block will sink”); (b) *has done* in the past (called an **observation**, e.g., “this block sank”); or (c) *is* (called a **conclusion**, e.g., “the density of this block is greater than 1.0 g/mL”). It is derived from, or explained by, the logical Application of the Rule to the particular circumstances described by the Premise.

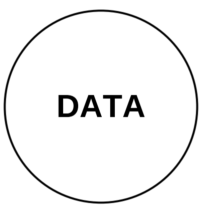
Also called: “claim” (Toulmin, et al., 1984); “claim” (McNeill & Krajcik, in press)

Inputs



A **PREMISE** is a statement describing the relevant characteristics or properties of the object about which the Claim is made. The Premise is the “given” information from whence the Claim is derived upon the Application of the Rule.

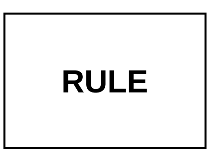
Also called: “data” (Toulmin, et al., 1984); “explanation” (Duschl, draft); “evidence” (McNeill & Krajcik, in press)



A **DATUM** is a statement describing the outcome of a single **specific** experiment. Related Data are summarized during Analysis to provide Evidence.

Also called: “set of collected data” (Duschl, draft)

Components



A **RULE** is a statement describing a **generalized** relationship, principle, or law. The Rule is general in the sense that it is defined in terms of and is expected to hold in contexts and circumstances **not previously observed**. The Rule is an Interpretation and generalization of the Evidence that is applied to the Premise in order to arrive at or explain the Claim.

Also called: “warrant” (Toulmin, et al., 1984); “patterns & models” (Duschl, draft); “reasoning” (McNeill & Krajcik, in press)

EVIDENCE

A piece of **EVIDENCE** is a statement summarizing a related set of Data. Because the Evidence is specific to the experimental context in which the Data were collected, it describes a **contextualized** relationship or finding, rather than a general principle or law. The Evidence is a summary resulting from an Analysis of the Data that can be generalized during Interpretation to form a Rule.

Also called: “backing” (Toulmin, et al., 1984); “evidence” (Duschl, draft)

Processes

APPLICATION

The **APPLICATION** is the process by which the Rule is applied to the specific circumstances described by the Premise. It establishes the probability or necessity of the Claim, often by logical deduction in simple cases. More complex cases with multiple Premises, Rules, and possible Claims may require systems analysis.

Also called: “qualifier” (Toulmin, et al., 1984); “Transformation 3” (Duschl, draft)

INTERPRETATION

The **INTERPRETATION** is the process by which the Evidence is generalized to create a Rule. In situations with multiple pieces of Evidence, Interpretation also involves the comparison and integration or synthesis of the Evidence. A common error of Interpretation is over-generalization.

Also called: “Transformation 2” (Duschl, draft)

ANALYSIS

The **ANALYSIS** is the process by which multiple pieces of Data are summarized to create a piece of Evidence, often by interpolation or model-fitting.

Also called: “Transformation 1” (Duschl, draft)

Figure 3. Description of the Parts of the Using Evidence Framework

The Using Evidence Framework is intended to help researchers and practitioners identify the components and processes of using evidence that are present in student work and discourse. Different aspects of the UE Framework can be selected as a focus for assessment and subsequent interpretation, for example, the number and structure of the components that are present, the conceptual sophistication of the Rule, and/or the type of Evidence that is brought to bear. In the papers that follow, we have chosen different aspects to focus on within the common Framework to illustrate some of the possibilities that are available.

At this point, it is appropriate to emphasize what the Using Evidence Framework is intended to model and what it is not. The framework is simply intended to model the use of evidence in scientific argumentation by both novices and experts. What the framework **does not** model is everything that teachers and students might say or write in a science classroom, general argumentation (although the UE Framework shares deep connections with models of general argumentation), or extended processes of logic and deduction, beyond the single application step.

In addition, while we believe that what a student says or writes during scientific argumentation can be mapped to some or all of the framework, we **do not believe** that every instance of scientific argumentation can be fully mapped to all of the framework. In fact, we believe that in the vast majority of cases what a student says or writes will only map to a portion of the framework. These beliefs are discussed in the *Mapping Talk and Writing to the Framework* section of this paper and will be illustrated in subsequent use of the framework to design assessment items (Nagashima et al., 2008).

Mapping Individual Statements to the Using Evidence Framework

In general, we assert that an individual statement made by a student can be mapped to one or more of either the Claim, Premise, Rule, Evidence, or Data parts of the framework. The classification of an individual statement cannot be made by semantic or syntactic analysis of the isolated statement. Instead, we believe that it is necessary to consider the location and purpose of the statement within the context of the entire argument, which may not be manifested, in a brief response.

For example, the statement “this block has a density greater than 1.0 g/mL…” could either be:

- a **Premise**, linked to the Claim “…therefore it will sink”

or

- a **Claim**, linked to the Premise “…because it sank”

In contrast, the three processes (Analysis, Interpretation, and Application) are often revealed by single words like “so” or “therefore” that link statements. Sometimes, however, these linking words are missing and the process must be inferred. For example, if a student states a Rule and makes a Claim, we can infer that the student engaged in the process of Application, but was not explicit about stating it.

In the example shown in Figure 4, the CLAIM “this metal cube will sink in a glass of water” is based on the APPLICATION (“therefore”) of the Rule “heavy things sink” to the PREMISE “this metal cube is heavy.” In turn, the RULE is based on two DATA points (DATUM1--“my heavy marble sank when I dropped it in the bathtub;” DATUM 2--“my heavy toy car sank when I dropped it in the bathtub”), which have undergone ANALYSIS (“therefore”) to provide the EVIDENCE “heavy toys I’ve dropped in my bathtub have sunk” leading to the INTERPRETATION (“therefore”) of the RULE.

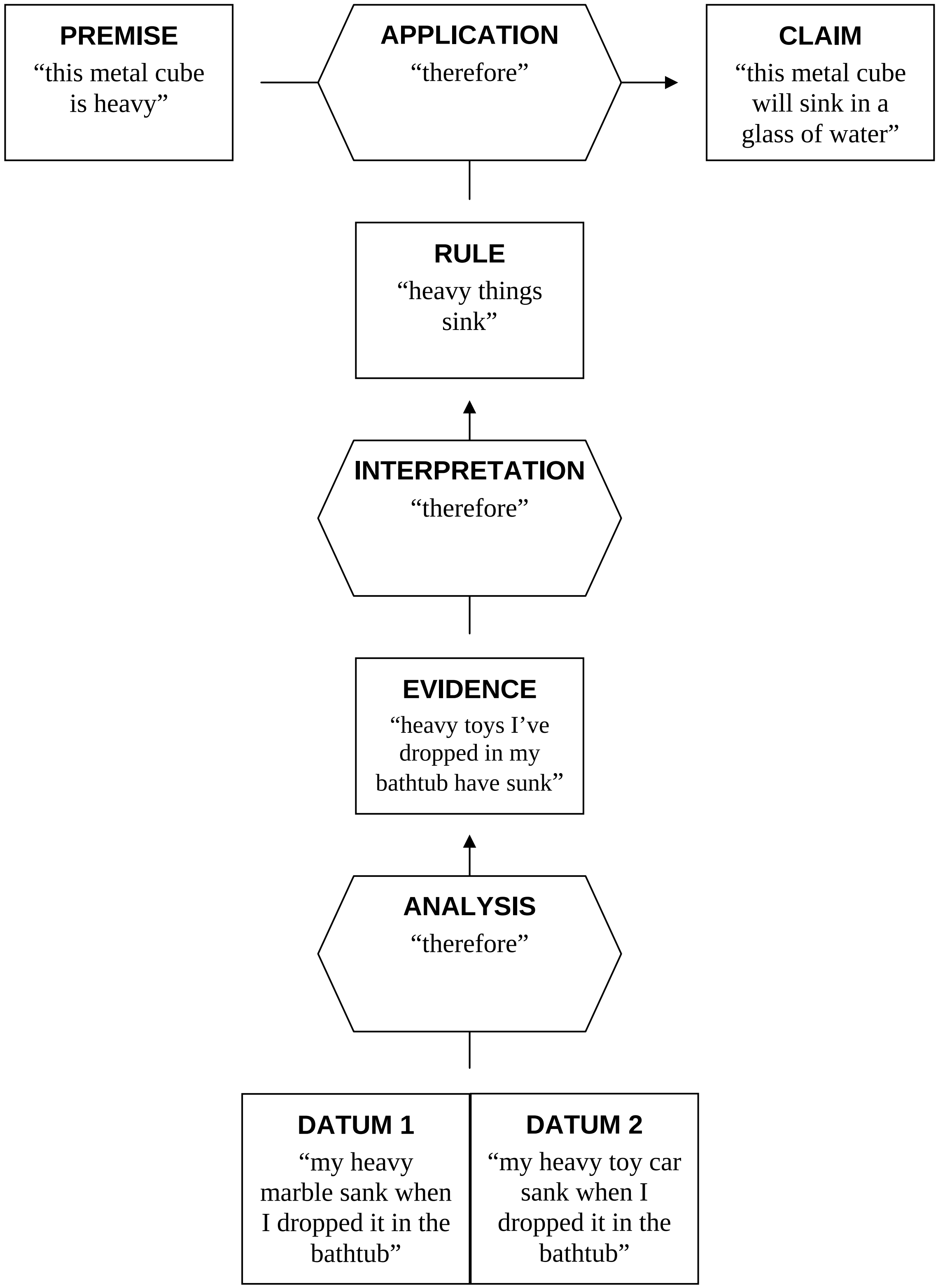


Figure 4. Simple Example of Mapping a Student Argument to the Framework

Mapping Individual Arguments to the Using Evidence Framework

We believe that sufficiently advanced students, with scaffolding, can produce all five statements (Claim, Premise, Rule, Evidence, and Data) for any given scientific argument. That is to say, we believe that all parts of the framework are always potentially observable. An example of a fully-mapped argument is shown in Figure 4. However, in the majority of cases, students will make statements that can be mapped only to *parts* of the framework. We propose that there is a hierarchy of sophistication in the levels of argumentation that a student might go through as they learn to apply all the parts of reasoning with evidence. This hierarchy is illustrated in Figure 5 in the sequence of diagrams labeled 1-4.

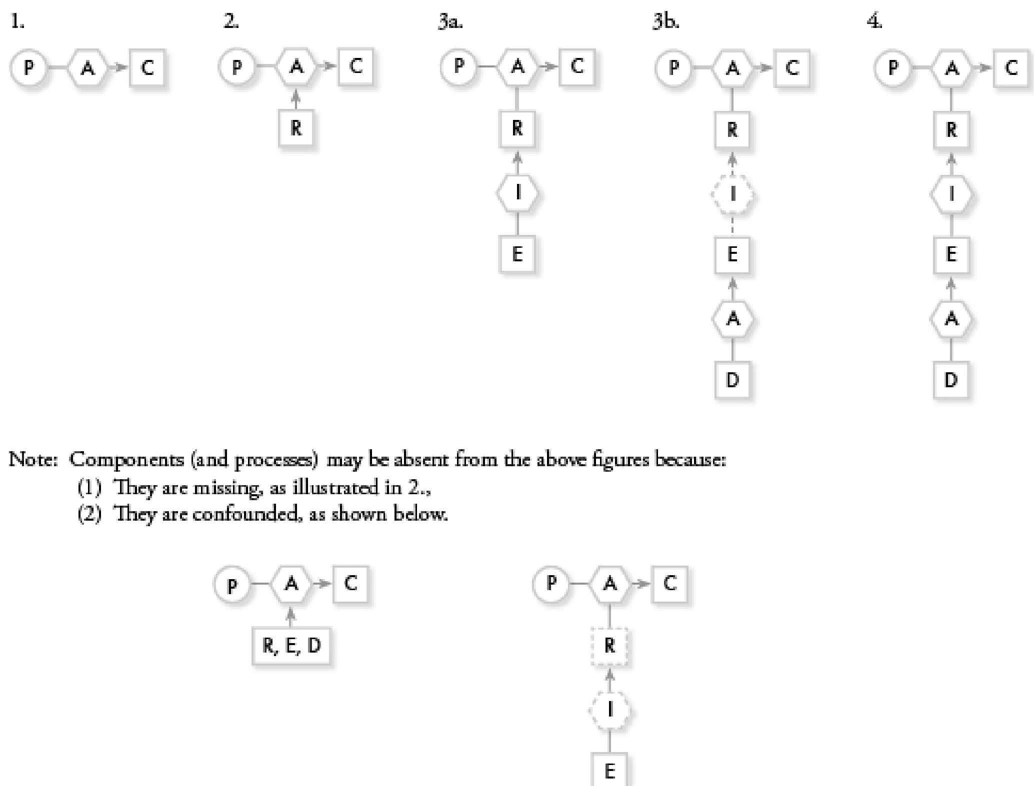


Figure 5. Examples of how Student Statements During Reasoning about Evidence Might Map to Parts of the Framework.

The diagram labeled 1 in Figure 5 shows an example of how a student statement might only apply a premise to make a claim. Relating this back to the example argument shown in Figure 4, the student might only say, “This metal cube is heavy, therefore, it will sink in water.” Part 2 of Figure 5 shows a more sophisticated argument that brings in a rule. Staying with the example in Figure 4, a student at this stage of reasoning might say, “This metal cube is heavy and I know that heavy things sink, therefore, it will sink in water.” Parts 3a and 3b show two examples of the next level of sophistication in an argument. In part 3a, the student brings evidence into their argument, although it is not linked to data, and may make a statement such as, “Heavy toys I’ve dropped in my bathtub have sunk so I think that heavy things sink and this metal cube is heavy, therefore, it will sink in water.” A similar level of reasoning is represented in part 3b where a student does refer to data, makes an analysis and produces evidence from it. They also start with a premise, apply a rule and make a claim, but they do not link the upper and lower sections of the framework because they do not interpret the data to produce the rule. For example, a student might say that “My heavy marble and my heavy toy car both sank when I dropped them in the bathtub, so all the heavy toys I’ve dropped in my bathtub have sunk. This metal cube is heavy and heavy things sink, therefore, it will sink in water.” Part 4 of Figure 5 represents the fully developed argument that employs all parts of the framework, as shown in the example of Figure 4.

We posit that there are reasons why students may not be able to produce a fully-fledged argument that employs all parts of the framework. First, they may not be intellectually capable of doing so because they have not reached an age where they have developed the ability to conceptualize the different parts of the framework. Second, some students may be developmentally capable of conceptualizing the differences between the parts of the framework, but still do not do so because they have not learned these differences. That is to say, students will likely need to receive instruction on the differences between the parts of the framework before they can produce all five statements for a given scientific argument. Third, students who have learned the framework may still not produce all five statements for rhetorical reasons, unless they are sufficiently scaffolded or prompted. Depending on the context, when asked to “justify” a Claim, some students may cite the Premise, other students may cite the Rule, and still others may cite examples of Data or Evidence. Even professional scientists rarely, if ever, appeal all the way back to individual pieces of Data collected in historical experiments when asked to support a Claim, despite being able to produce such Data if needed.

Currently, we do not know the extent to which it is possible to distinguish between these three interpretations. At a minimum, we recommend sufficient scaffolding so that the third interpretation can be eliminated.

Mapping conflicting arguments and counterevidence

The framework also provides a way of representing scientific argumentation between two or more people. When two people are arguing, or when one person is weighing counterevidence, this can be represented by using one diagram for each argument, as illustrated in Figure 6.

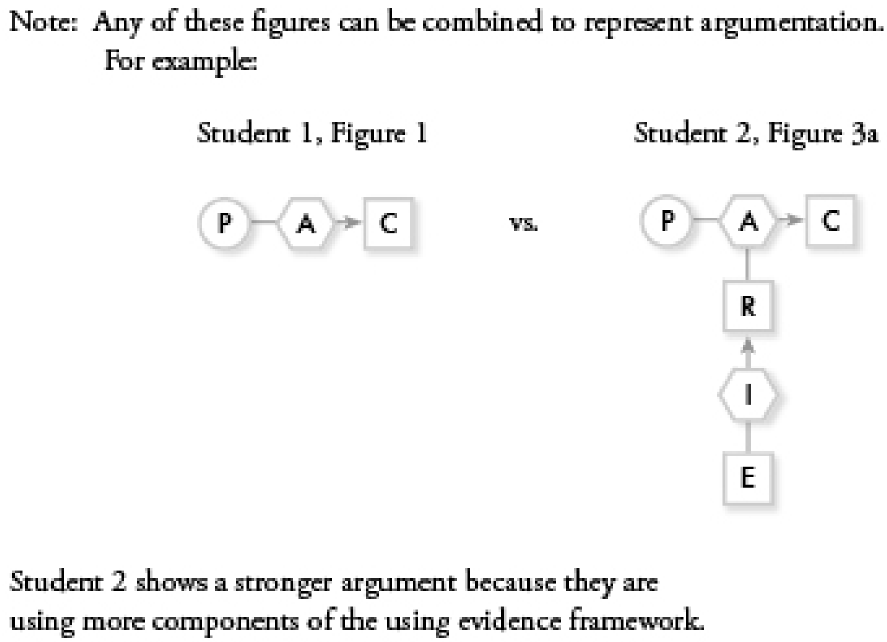


Figure 6. Example of a Two-sided Scientific Argument

Figure 6 shows how a scientific argument between student 1 and student 2 might be mapped to the framework as a way of judging which argument has more intellectual and scientific "weight." Student 1 is only arguing at level 1 (see Figure 5) but student 2 is arguing at level 3a, so student 2's argument is scientifically stronger.

Application of the Using Evidence Framework in Assessing Talk and Writing

In applying the Framework to students' talk or writing as they use evidence, there are some things to clarify about how to assess the outputs (Claim), inputs (Premise or Data), and components (Rule or Evidence). First, in assessing the Claim, we do not independently assess the quality of the claim apart from the quality of the components and processes making up the argument (see below). If the Analysis, Interpretation, and Application are all valid, then the

Claim will also be valid. If the Rule is conceptually sophisticated and precise, and the Evidence is precise and reliable, then the Claim will also be accurate and reliable.

Also, we do not assess the quality of the Premise or the Data, as these are taken as the "givens" within the context of the argument. Whether or not the Premise has been established correctly or the Data have been properly collected are important issues, but outside the scope of the UE Framework *per se*.

Furthermore, the Rule can be assessed in terms of its Conceptual Sophistication and its Precision. Conceptual Sophistication refers to the quality and complexity of the concepts that the Rule implicates, ranging from misconceptions to normative scientific conceptions. Precision refers to the degree of specificity in how the Rule is phrased, ranging from ambiguous statements to quantified statements with appropriate units. This is fully explained by Nagashima et al., (2008) in their description of how the UE Framework was used to construct assessment items designed to measure student's use of evidence and how the scoring used the two variables of Conceptual Sophistication and Precision.

Similarly, the Evidence part can be assessed in terms of its Precision and its Reliability. Precision is characterized in the same way as for the Rule (see above). Reliability refers to the quality of the source and the quality and quantity of the Data that makes up the Evidence, ranging from made up examples to controlled experiments with multiple trials. This use of this variable is described by Nagashima et al. (2008). Finally, the processes of Application, Interpretation and Analysis can be assessed in terms of their Validity. Validity refers to the quality of the reasoning linking one component to another, ranging from no link to valid logical connections. Again, how this can be done is described by Nagashima et al. (2008).

Conclusions

It is hoped that this analytic framework of the use of evidence in scientific reasoning will form a robust platform for developing research instruments that can detect and measure students' and their teachers' use of evidence in science teaching. The utility of the framework was tested by a collaborative team of researchers from research institutions and universities in the United States and in Germany in a series of three studies. In the first study, we describe how the framework was used to develop test items that were designed to probe students' use of evidence and the findings when the items were administered to middle school students in the US (Nagashima et al., 2008). In a second study, the framework was used in developing the *Reasoning in Science Classroom Discourse* instrument that was created to measure the quality of student reasoning in whole-class discussions and to capture teachers' and students' co-constructed reasoning about scientific phenomena (Furtak et al., 2008). In a third study, *Reasoning in Science Classroom Discourse* was applied to two datasets of elementary science instruction on the topic of floating and sinking (Hardy et al., 2008). Both datasets are based on instructional variations designed to promote students' understanding of scientific concepts, one involving different degrees of scaffolding, the other involving prior instruction on nature-of-science constructs.

We plan to develop other applications of the UE Framework and we hope that other researchers, test developers and curriculum developers will also apply the framework in the creation of research instruments, assessments and instructional materials that test the framework's utility further. Over time we hope that this will lead to improvements in teaching students the key skills of using evidence in scientific reasoning and thereby promote growth in students' understanding of science concepts.

REFERENCES

American Association for the Advancement of Science. (1993). *Benchmarks for Science Literacy*. New York: Oxford University Press.

~~Adams, R. J., Wilson, M., & Wang, W. (1997). The multidimensional random coefficients multinomial logit model. Applied Psychological Measurement, 21, 1-23.~~ Bell, P. & Linn, M.C. (2000). Scientific arguments as learning artifacts: Designing for learning from the web with KIE. *International Journal of Science Education*, *22*(8), 797-817.

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the Norms of Scientific Argumentation in Classrooms. *Science Education, 84*, 287-312.

Duschl, R. A. (2003). Assessment of Inquiry. In J. M. Atkin & J. Coffey (Eds.), *Everyday Assessment in the Science Classroom* (pp. 41-59). Arlington, VA: NSTA Press.

~~Duschl, R. A., & Gitomer, D. H. (1997). Strategies and Challenges to Changing the Focus of Assessment and Instruction in Science Classrooms. *Educational Assessment, 4*(1), 37-73.~~

Duschl, R.A., (draft). The HS Lab Experience: Reconsidering the Role of Evidence, Explanation and the Language of Science.

Erduran, S., Simon, S., & Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. Science Education, 88, 915--933.

Furtak, E.M., Hardy, I., Beinbrech, T., Shavelson, R.J. and Shemwell, J.T. (2008). A Framework for Analyzing Reasoning in Science Classroom Discourse. Paper to be presented at the Annual Meeting of the American Educational Research Association, New York, March 2008.

Hardy, I. ,Kloetzer, B., Möller, K., and Sodian, B. (2008). The Analysis of Classroom Discourse: Elementary School Science Curricula Advancing Reasoning with Evidence. Paper to be presented at the Annual Meeting of the American Educational Research Association, New York, March 2008.

Hempel, C. G. (1966). *Philosophy of Natural Science*. Englewood Cliffs, N.J.: Prentice-Hall.

Jimenez-Alexandre, M. P., Rodriguez, A. B., & Duschl, R. A. (2000). "Doing the Lesson" or "Doing Science": Argument in High School Genetics. *Science Education, 84*, 757-792.

Kelly, G. J., & Bazerman, C. (2003). How Students Argue Scientific Claims: A Rhetorical-Semantic Analysis. *Applied Linguistics, 24*(1), 28-55.

Kuhn, D. (1993). Science as Argument: Implications for Teaching and Learning Scientific

Thinking. *Science Education*, *77*(3), 319-337.

~~Kuhn, D. (1993). Science as Argument: Implications for Teaching and Learning Scientific Thinking. *Science Education*, *77*(3), 319-337.~~

McNeill, K.L. & Krajcik, J. (in press). Assessing middle school students' content knowledge and reasoning through written scientific explanations. In Coffey, J., Douglas, R., & Binder, W. (Eds), *Science Assessment: Research and Practical Approaches*. Arlington, VA: National Science Teachers Association Press.

McNeill, K. L., Lizotte, D. J., Krajcik, J., & Marx, R. W. (2006). Supporting Students' Construction of Scientific Explanations by Fading Scaffolds in Instructional Materials. *The Journal of the Learning Sciences*, *15*(2), 153-191.

~~National Research Council (2001). Knowing what students know: The science and design of educational assessment. Washington, D.C.: National Academy Press.~~

Nagashima, S.O., Brown, N.J.S., Fu, A., Timms, M.J. and Wilson, M.R. (2008). *A Framework for Analyzing Reasoning in Written Assessments*. Paper to be presented at the Annual Meeting of the American Educational Research Association, New York, March 2008.

National Research Council. (1996). *National Science Education Standards*. Washington, D.C.: National Academy Press.

National Research Council (2001). Knowing what students know: The science and design of educational assessment. Washington, D.C.: National Academy Press.

National Research Council. (2007). *Taking Science to School: Learning and Teaching Science in Grades K-8.* Washington, D.C.: National Academies Press. ~~Newton, D. P./Newton, L. D. (2000): Do Teachers support Causal Understanding through their Discourse when Teaching Primary Science? In: British Educational Research Journal. Vo. 26, No. 5. 599-613.~~

Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the Quality of Argumentation in School Science. *Journal of Research in Science Teaching, 41*(10), 994-1020.

Sandoval, W. A. (2003). Conceptual and Epistemic Aspects of Students' Scientific Explanations. *The Journal of the Learning Sciences, 12*(1), 5-51.

Sandoval, W. A., & Millwood, K. A. (2005). The Quality of Students' Use of Evidence in Written Scientific Explanations. *Cognition and Instruction*, *23*(1), 23-55.

Schwab, J. J. (1962). The Concept of the Structure of a Discipline. *The Educational Record, 43*(3), 197-205.

Simon, S., Erduran, S., & Osborne, J. (2006). Learning to Teach Argumentation: Research and development in the science classroom. *International Journal of Science Education, 28*(2-3), 235-260.

~~Sodian, B., Jonen, A., Thoermer, C. & Kircher, E. (2006). Die Natur der Naturwissenschaften verstehen – Implementierung wissenschaftstheoretischen Unterrichts in der Grundschule. In M. Prenzel, & L. Allolio-Näcke (Hrsg.). *Untersuchungen zur Bildungsqualität von Schule*. Münster: Waxmann.~~

Toulmin, S. E. (1958/2003). *The Uses of Argument*. Cambridge: Cambridge University Press.

Toulmin, Rieke, & Janik (1984). An introduction to reasoning.

Tytler, R., & Peterson, S. (2005). A Longitudinal Study of Children's Developing Knowledge and Reasoning in Science. *Research in Science Education, 35,* 63-98.