ADAPTING MODEL ANALYSIS FOR THE STUDY OF PROOF SCHEMES

Todd CadwalladerOlsker, California State University, Fullerton
David Miller, West Virginia University
Kelly Hartmann, California State University, Fullerton

This theoretical paper describes model analysis, a quantitative research method used in physics education research, and our adaptation of this method to the study of proof schemes in a transition to proof course. Model analysis accounts for the fact that students may hold more than one idea or conception at a time, and may use different ideas and concepts in response to different situations. Model analysis is uniquely suited to study students' proof schemes, as students often hold multiple, sometimes conflicting proof schemes, which they may use at different times. Model analysis treats each student's complete set of responses as a data point, rather than treating each individual response as a separate data point. Thus, model analysis can capture information on the self-consistency of a student's responses. We collected data from a Transition to Proof course at the beginning and at the end of the Fall 2012 semester. We then analyzed this data using traditional descriptive statistics as well as model analysis. We find that model analysis offers significant insights not offered by traditional analysis.

Key words: Mathematical Proofs, Proof Schemes, Model Analysis, Transition to Proofs

Introduction and Background

The idea of proof is central to mathematics, but the ability to write (or even to read) proofs is notoriously difficult for students to develop. One barrier to student success in proof writing is a poor concept of what makes a mathematical proof. Several researchers have developed taxonomies to describe the various notions of proof held by students; perhaps the most extensive is Harel and Sowder's (1998, 2007) taxonomy of proof schemes. Proof schemes describe the type of argument an individual (or group) finds to be a convincing mathematical proof. This taxonomy has been successful in helping researchers to understand one source of students' difficulty with mathematical proof: when students have inadequate proof schemes, they have difficulty in writing, or even understanding, mathematical proof (Recio & Godino, 2001; Housman & Porter, 2003; Zazkis & Liljedahl, 2004).

While current research has successfully used the taxonomy of proof schemes to understand some of the difficulties that individuals have in writing proofs, Harel and Sowder (2007) have asserted that proof schemes can develop at the level of a group or community; for instance, the community of learners in a given classroom. To our knowledge, no current research has examined proof schemes at the community level.

In this paper we will discuss how we adapted the method of model analysis, a technique pioneered by physics education researchers Bao and Redish (2001), to analyze proof schemes. Model analysis treats each student's complete set of responses as a data point, rather than treating each individual response as a separate data point. Thus, model analysis can capture information on the self-consistency of a student's responses. This allows the data to be analyzed with the understanding that students do not always use the same conceptions in response to every situation. This point is particularly important for analyzing proof schemes, as students often use different proof schemes to understand or write proofs in different contexts.

Our research question is the following:

What insights can model analysis provide into community-level proof schemes that traditional quantitative analysis cannot?

Review of Relevant Literature

We primarily draw from two pieces of research: Harel and Sowder's taxonomy of proof schemes, and Bao and Redish's method of model analysis.

Proof Schemes

Harel and Sowder have developed a framework for describing students' overall conception of what justifies a mathematical proof, which they refer to as a proof scheme. The original framework (Harel & Sowder, 1998) is one of several understandings of proof discussed by Balache (2002), and is perhaps the most extensive of such taxonomies. This framework of proof schemes has been used in several studies (Recio & Godino, 2001; Housman & Porter, 2003; Zazkis & Liljedahl, 2004) as a taxonomy for students' conceptions of their justification for proof. Harel and Sowder define proof scheme with the following: "A person's (or community's) proof scheme consists of what constitutes ascertaining and persuading for that person (or community)" (Harel & Sowder, 2007). Harel and Sowder's taxonomy consists of seven major types of proof scheme, organized into three broader categories: external conviction, empirical, and deductive. External conviction proof schemes are possessed by students who are convinced a theorem is true by external forces; empirical proof schemes describe the proof schemes of students who are convinced by evidence, rather than logical reasoning; deductive proof schemes construct and validate theorems by means of logical deductions. An individual's proof scheme may not consistently fit into a single category. Rather, the category of proof scheme can depend on context, and different contexts may activate different categories of proof scheme.

Model Analysis

The proposed research will make use of model analysis, a quantitative analysis technique developed by Bao and Redish (1999; 2001) for analyzing students' concepts of force in physics education research. According to Bao and Redish (2001), "The method is especially valuable in cases where qualitative research has documented that students enter a class with a small number of strong naive conceptions that conflict with or encourage misinterpretations of the scientific view." Michael Oehrtman (2006) has successfully used model analysis to analyze students' concepts of functions in first-year calculus courses.

The term model, as used by Bao and Redish, refers to a mental model of a particular concept, including ideas, conceptions, and beliefs about the concept. Proof schemes, then, can be thought of as a "model" of the concept of proof. Model analysis will be detailed in the next section, where we describe how it will be adapted for our project.

Bao and Redish describe the method of model analysis as consisting of five steps (2006). These steps, and how they were adapted for our project, are as follows:

> (i) Through systematic research and detailed student interviews, common student models are identified and validated so that these models are reliable for a population of students with a similar background. (Bao & Redish, 2006)

This step describes what Harel and Sowder accomplished when creating their taxonomy of proof schemes (Harel & Sowder, 1998). These categories were used as the "models" described by Bao and Redish.

> (ii) This knowledge is then used in the design of a multiple-choice instrument. The distracters are designed to activate the common student models, and the effectiveness of the questions is validated through research. (Bao & Redish, 2006)

For this step, we use the questionnaire developed by Stylianou and Blanton (in press). The effectiveness of these questions has been validated by their research. The questionnaire present four theorems, and for each theorem, asks students to choose from four different arguments which they believe to be the closest to their own approach, which they consider to be the most rigorous, and which they consider the most explanatory. Stylianou and Blanton categorize each of these responses with a proof scheme (deductive or empirical) and a style (symbolic or narrative for deductive proofs, visual or numeric for empirical proofs) for a total of four types. We have modified the questionnaire slightly by allowing a response of “None of the above" to some questions, but this alteration should not invalidate the effectiveness of the questions.

> (iii) One then characterizes a student's responses with a vector in a linear “model space" representing the (square roots of the) probabilities that the student will apply the different common models. (Bao & Redish, 2006)

The model space described in this step is represented mathematically by a linear vector space, where each common model is represented by an element of an orthonormal basis. That is, each of the categories of proof scheme (empirical and deductive) will be assigned a dimension in the vector space. We also assigned (as do Bao and Redish) a third dimension to a “null" model, considered to be activated when students choose a response of “none of the above." For each student, we created a vector inside this model space that represents the student's responses to the questionnaire. Each entry in the vector is meant to represent the probability with which the student uses the associated category of proof scheme to respond to similar types of questions. Of course, these probabilities can only be approximated by the student's responses to the questionnaire.

> (iv) The individual student model states are used to create a “density matrix," which is then summed over the class. The off-diagonal elements of this matrix retain information about the confusions (probabilities of using different models) of individual students. (Bao & Redish, 2006)

For each model state vector, a density matrix is created by taking the outer product of the model state vector with itself. The diagonal entries of the density matrix are simply the probabilities calculated in step (iii). The off-diagonal entries are non-zero only when a student has responses from more than one category of proof scheme. Thus, when a student is inconsistent, those inconsistencies are preserved by the off-diagonal entries of the matrix.

To study a large number of data points, the density matrices will be averaged together: that is, the entries in each position are added together and divided by the total number of data points. Bao and Redish refer to the resulting matrix as the class density matrix.

> (v) The eigenvalues and eigenvectors of the class density matrix give information not only how many students got correct answers, but about the level of confusion in the state of the class's knowledge. (Bao & Redish, 2006)

The class density matrix contains information on the students' responses to the questionnaire. An eigenvalue decomposition allows for trends in the data to be identified. In this way, the eigenvalue decomposition allows for information about the class as a whole to be extracted from the data. That is, the community-level proof scheme held by the class can be identified.

Data and Results

Our study collected data at the beginning (pre-instruction) and at the end (post-instruction) of the Fall 2012 semester in a Transition to Proofs class. We used a multiple choice questionnaire instrument developed by Stylianou and Blanton (in press) to collect data from students in order to identify the proof schemes held in this introductory proof-writing class. A total of 38 students participated in both the pre- and post-instruction surveys.

Tables 1 and 2, below, present the class density matrix and eigenvalues/vectors derived from model analysis. Table 1 compares the pre- and post-instruction results derived from the questions asking which proof is the most rigorous; Table 2 compares those derived from the questions asking for the most explanatory proof. Not shown here are the results from questions asking for the proof closest to students' approach. A discussion of each table follows.

Pre-instruction:
Class Density Matrix:

	N-A	E-V	E-N	D-N	D-S
None of the Above	0.0066	0.0000	0.0000	0.0066	0.0093
Empirical-Visual	0.0000	0.0921	0.0273	0.0449	0.1004
Empirical-Numeric	0.0000	0.0273	0.1645	0.0525	0.1219
Deductive-Narrative	0.0066	0.0449	0.0525	0.1645	0.2023
Deductive-Symbolic	0.0093	0.1004	0.1219	0.2023	0.5724

Eigenvalues/vectors:

Eigenvalues	0.0062	0.0705	0.1310	0.0820	0.7102
Associated Eigenvectors:					
None of the Above	0.9985	-0.0406	-0.0209	0.0261	0.0151
Empirical-Visual	0.0285	0.9444	-0.0057	0.2733	0.1803
Empirical-Numeric	0.0168	-0.0271	0.9689	-0.0467	0.2409
Deductive-Narrative	-0.0425	-0.3208	-0.0569	0.8712	0.3648
Deductive-Symbolic	-0.0099	-0.0523	-0.2399	-0.4044	0.8810

Post-instruction:
Class Density Matrix:

	N-A	E-V	E-N	D-N	D-S
None of the Above	0.0192	0.0128	0.0091	0.0128	0.0155
Empirical-Visual	0.0128	0.0449	0.0356	0.0245	0.0219
Empirical-Numeric	0.0091	0.0356	0.1154	0.0091	0.0175
Deductive-Narrative	0.0128	0.0245	0.0091	0.1603	0.2073
Deductive-Symbolic	0.0155	0.0219	0.0175	0.2073	0.6346

Eigenvalues/vectors:

Eigenvalues	0.0134	0.0294	0.1318	0.0847	0.7150
Associated Eigenvectors:					
None of the Above	0.8877	0.4356	0.1184	0.0863	0.0286
Empirical-Visual	-0.4531	0.7831	0.3966	0.1488	0.0458
Empirical-Numeric	0.0806	-0.3552	0.8906	-0.2699	0.0357
Deductive-Narrative	0.0097	-0.2595	0.1502	0.8865	0.3522
Deductive-Symbolic	-0.0117	0.0597	-0.1139	-0.3340	0.9337

Table 1: Class density matrices and eigenvalues/vectors derived for "most rigorous" questions (n=38)

Looking at the results from the questions asking for the most mathematically rigorous proof, we see that even pre-instruction, students tend to choose deductive-symbolic responses. The diagonal entries of the class density matrix indicate that approximately 57% of the responses are in the deductive-symbolic category, 16% are in each of the deductive-narrative and empirical-numeric category, and 9% in the empirical-visual category. All of these percentages are easily obtained with traditional statistical methods, as well. The off-diagonal entries show the level of student self-consistency, which cannot be easily obtained with traditional methods. Higher off-diagonal entries indicate that students tended to choose responses from both of those categories, and therefore lower levels of self-consistency. The pre-instruction class density matrix indicates relatively high off diagonal entries in the deductive-symbolic row, indicating that a fair number of students chose empirical-visual, empirical-numeric, and (especially) deductive-narrative responses in addition to the majority of deductive-symbolic responses. The post-instruction class density matrix, by contrast, indicates that the "overlap" between deductive-narrative and deductive-symbolic remains, whereas the overlap between deductive-symbolic and the empirical categories has all but disappeared.

The eigenvector analysis paints a similar picture: in both pre-and post-instruction results, there is a dominant eigenvector, associated with an eigenvalue much larger than any of the others. Bao and Redish (2006) indicate that an eigenvalue of 0.8 or greater indicates a strong primary eigenvalue; this eigenvalue does not quite meet this threshold. Nonetheless, this eigenvector indicates that the primary model state held by the class has a strong tendency toward the deductive-symbolic, a weaker tendency toward the deductive-narrative, and weak tendencies toward the other models. The dominant eigenvector the post-instruction results has very weak tendencies toward the empirical models.

Pre-instruction:
Class Density Matrix:

	N-A	E-V	E-N	D-N	D-S
None of the Above	0.0197	0.0159	0.0225	0.0132	0.0000
Empirical-Visual	0.0159	0.2500	0.1505	0.0871	0.0770
Empirical-Numeric	0.0225	0.1505	0.2368	0.0476	0.0300
Deductive-Narrative	0.0132	0.0871	0.0476	0.1645	0.1305
Deductive-Symbolic	0.0000	0.0770	0.0300	0.1305	0.3289

Eigenvalues/vectors:

Eigenvalues	0.0163	0.2945	0.0954	0.0772	0.5165
Associated Eigenvectors:					
None of the Above	0.9870	0.0612	-0.0146	0.1395	0.0481
Empirical-Visual	0.0212	0.4077	-0.4586	-0.5662	0.5500
Empirical-Numeric	-0.0973	0.5741	0.5931	0.3494	0.4325
Deductive-Narrative	-0.1146	-0.1931	-0.5428	0.6966	0.4121
Deductive-Symbolic	0.0519	-0.6806	0.3783	-0.2294	0.5817

Post-instruction:
Class Density Matrix:

	N-A	E-V	E-N	D-N	D-S
None of the Above	0.0064	0.0091	0.0064	0.0000	0.0000
Empirical-Visual	0.0091	0.2179	0.0750	0.0467	0.0605
Empirical-Numeric	0.0064	0.0750	0.1795	0.0586	0.0613
Deductive-Narrative	0.0000	0.0467	0.0586	0.2051	0.1104
Deductive-Symbolic	0.0000	0.0605	0.0613	0.1104	0.3654

Eigenvalues/vectors:

Eigenvalues	0.0059	0.2254	0.1148	0.1469	0.4814
Associated Eigenvectors:					
None of the Above	0.9988	0.0426	0.0090	-0.0197	0.0110
Empirical-Visual	-0.0379	0.6948	-0.4563	-0.4335	0.3460
Empirical-Numeric	-0.0274	0.4733	0.7975	0.1817	0.3259
Deductive-Narrative	0.0131	0.0393	-0.3854	0.8135	0.4336
Deductive-Symbolic	0.0070	-0.5384	0.0849	-0.3419	0.7655

Table 2: Class density matrices and eigenvalues/vectors derived for "most explanatory" questions (n=38)

The class density matrix for the "most explanatory" data shows a much more mixed picture. While the largest diagonal entry of the class density matrix is still in the deductive-symbolic category, the other diagonal entries are almost as large. However, the off-diagonal entries are slightly smaller overall than they were for the "most rigorous" results, indicating a slightly higher level of self-consistency. The largest overlaps are between the two empirical categories and between the two deductive categories.

The eigenvalue decompositions for the "most explanatory" class density matrices gives much weaker primary eigenvalues. The associated eigenvector shows a mix of model states, and is not so different from that associated with the next highest eigenvalue. Bao and Redish suggest that this kind of result indicates that the model states of the students are not very orthogonal, and that the eigenvalue decomposition is less valuable in this case.

Conclusion

The main claim of this paper is that the technique of model analysis can tell researchers about the self-consistency of student responses to Stylianou and Blanton's proof scheme items. Our results from the "most rigorous" questions are encouraging: students entered the Transition to Proof class with a fairly self-consistent deductive model, and ended the course with an even stronger, more self-consistent, deductive model. The results from the "most explanatory" questions are less encouraging; even though the post-instruction results skew more toward the deductive models, there is not a consensus among the students that deductive models are the most explanatory.

It is important to note that our data set is very limited, and that our results are based on a small sample size. However, our data does show that model analysis can differentiate between self-consistent data and non-self-consistent data. We believe that model analysis provides a valuable perspective on data collected using Stylianou and Blanton's survey. We currently plan to apply this method to a more robust data set collected by Stylianou and Blanton.

References

Balacheff, N. (2002). A researcher epistemology: A deadlock for educational research on proof. In Proceedings of the 2002 international conference on mathematics: Understanding proving and proving to understand (pp. 23-44). Department of Mathematics, National Taiwan Normal University.

Bao, L. (1999). Dynamics of student modeling: A theory, algorithms, and application to quantum mechanics. University of Maryland, College Park: Doctoral dissertation.

Bao, L., & Redish, E. F. (2001). Concentration analysis: A quantitative assessment of student states. American Journal of Physics, 69 (S1), S45-S53.

Bao, L., & Redish, E. F. (2006). Model analysis: Representing and assessing the dynamics of student learning. Physical Review Special Topics { Physics Education Research, 2 (010103).

Harel, G., & Sowder, L. (1998). Students proof schemes: Results from exploratory studies. In A. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), Research in collegiate mathematics education III (pp. 234-282). American Mathematical Society: Washington, D.C.

Harel, G., & Sowder, L. (2007). Toward a comprehensive perspective on proof. In F. Lester (Ed.), Second handbook of research on mathematics teaching and learning. National Council of Teachers of Mathematics.

Housman, D., & Porter, M. (2003). Proof schemes and learning strategies of above-average mathematics students. Educational Studies in Mathematics, 53 , 139-158.

Oehrtman, M. (2006). Model analysis in undergraduate mathematics education research. In Abstracts of the ninth conference conference on research in undergraduate mathematics education.

Recio, A. M., & Godino, J. D. (2001). Institutional and personal meanings of proof. Educational Studies in Mathematics, 48 (1), 83-99.

Stylianou, D. A., Blanton, M. L., & Rotou, O. (In Press). Undergraduate students' understanding of proof: Relationships between proof conceptions, beleifs, and classroom experiences with learning proof. In D. Holton & P. Thompson (Eds.), Research in collegiate mathematics education. American Mathematical Society: Washington, D.C.

Zazkis, R., & Liljedahl, P. (2004). Understanding primes: The role of representation. Journal for Research in Mathematics Education, 35 (3), 164-186.