AN AGENT-BASED APPROACH FOR EDUCATION MODELING

Joanne Wendelberger, Benjamin Sims (Statistical Sciences)  
Stephan Eidenbenz, Patrick Kelly (Information Sciences)  
Steven Stringer (Technology Transfer)  
Los Alamos National Laboratory

Michael Strong  
University of California Santa Cruz

Introduction

The research literature on the factors that predict mathematics performance among secondary school students is typically subdivided into distinct areas such as instruction, motivation, academic achievement, gender, ethnicity, home background, and so on. With the exception of a small number of multilevel studies, most researchers tend to focus on variables that belong to one set of predictive components to the exclusion of others.

This project examines an approach to address the research question: What are the school-related predictors of mathematics success of public school students in 8th grade and beyond? The approach uses an agent-based model (ABM) for modeling and simulation of an educational system. The main task of this project was to initiate the development and implementation of an ABM to assess the feasibility of using the ABM approach to address issues such as the research question above. This project involves a scientific approach to predictive modeling that incorporates capabilities in mathematics, computer science, and domain expertise.

Our study uses a large longitudinal school district database, which includes a wide range of potentially predictive variables, as a vehicle for applying an agent-based modeling approach in order to address the complexity of the problem of student math achievement. This approach, together with the application of advanced visualization techniques, is designed as a proof-of-concept study that will enable us to examine and demonstrate how a variety of student, teacher, and environmental parameters may interact to influence a learner’s mathematics performance in the upper school grades. The ultimate goal is to develop a model of mathematics achievement that can inform school administrators in their efforts to develop the most effective educational interventions for their students. The project will have reached its goal if:

• An agent-based model is produced that demonstrates its ability to simulate behavior arising from student, teacher, and environmental interactions.
• Statistical analysis is conducted on the school district dataset to inform and calibrate the agent-based model to produce simulated data that is consistent with observations from actual school district data.
• Simulations and what-if scenarios are run to test the interactions between and among parameters under a variety of circumstances and interventions.
The resulting data can be presented using visualization techniques that convey the information in a compelling manner to potential end users. Educators and policy makers confirm that this approach holds promise for understanding the interactions among key factors relevant to K-12 student success in mathematics.

In the sections that follow, background and development of a conceptual model for an educational system is discussed, followed by a description of the implementation of an actual simplified prototype model and the associated data analysis and calibration procedures used to incorporate actual data from a school district into the modeling process. Together, the conceptual framework, agent-based simulation modeling, and statistical methodology provide an approach for examining how students move through an educational system and for studying the impact of different factors on the resulting system behavior.

Background

The work conducted at Los Alamos National Laboratory includes development of a conceptual framework for describing an educational system, construction of an agent-based model to represent an educational system, and incorporation of data from the San Jose Unified School District into the modeling process. A preliminary model has been developed to demonstrate the agent-based modeling approach for a simplified educational system. The model tracks the behavior of students as they move through classrooms in a school over time, interacting with teachers and other students. The preliminary model incorporates the following core elements:

- model entities and their attributes (students, teachers, classrooms, etc.),
- a structure to guide entity interactions (school system, schools, classes),
- a process for assigning students and teachers to school systems, schools, classrooms, and classes,
- functions for assigning scores/grades to students (based on student attributes, interactions with teachers and other students, and previous scores/grades in transcript), and
- methods for capturing and displaying model outputs.

In conjunction with the modeling effort, variables in the San Jose Unified School District database were examined to identify and characterize variables that will inform the agent-based model. Initial focus is on information from California state test scores that is directly tied to mathematics performance required for participation in algebra. This information was used to inform and calibrate the model to generate behavior that is consistent with the behavior of actual students.

This project brings together educational expertise, state-of-the-art modeling and simulation techniques, and statistical characterization and calibration methodology to develop a representation of an educational system that incorporates individual behavior as part of a dynamic structure of interacting components. This effort leverages technology developed at Los Alamos National Laboratory for modeling transportation systems and other infrastructure systems involving complex human interactions. The approach goes beyond routine data analysis and instead seeks to build a model that is grounded in educational theory and informed by statistical information from the educational database developed by the San Jose Unified School District. Demonstration of the use of agent-based modeling in this manner provides an approach for examination of an educational system and quantitative evaluation of factors impacting educational progress in a rigorous, scientific manner.
A Conceptual Model of the Education Process

Conceptual Model Development Process

Development of a conceptual model is part of a larger process for ensuring the development of a sound computational model (Figure 1). Our approach draws on the model evaluation approach developed by Sargent (2005) and expanded by McNamara et al. (2008). The conceptual model is developed primarily with input from education experts and captures their understanding of the education domain, along with ideas derived from initial data analysis and computational model development. Ongoing efforts in socio-technical integration ensure that the conceptual model, the statistical analysis, and the computational model remain consistent with each other, and that the education experts, computer scientists, and statisticians on the team are in frequent communication. In the later stages of the project, statistical analysis serves to validate the computational model, and education experts evaluate its credibility as a representation of the education process. This socio-technical integration was coordinated by a sociologist knowledgeable in both expert knowledge capture and development and validation of social models. The integrated nature of this process ensures that model credibility is established from the early stages of the project, well before any formal validation process takes place. The development of the conceptual model plays a crucial role in this early establishment of model credibility.

![Figure 1. Model integration and evaluation process](image)

The initial stages of conceptual model development were 1) expert knowledge capture and 2) review of existing theories of the education process. Initial conversations about education concepts culminated in a 3-hour formal knowledge capture session with the education professionals participating in the project. This session identified 47 potential factors contributing to student success or failure, including variables related to students (both demographics and personal abilities), teachers (ability and training), classroom environment (prerequisite knowledge, aggregate student characteristics, etc.), schools (resources, etc.),
and district and state policies (curriculum, teaching standards). The full list of variables elicited in this session is provided in Appendix 1.

An in-depth literature review of research on the education process was conducted, and identified a number of relevant theoretical constructs and variables that have been connected to student achievement (Appendix 2). Since the goal of the project is ultimately to provide practical information to school, district, and policy personnel in order to help transform education, one way of organizing the variables is according to their susceptibility to change (their mutability). A number of variables describe intrinsic student characteristics at the low end of the mutability scale. Most prominent of these are gender and race, to which we may wish to add aptitude, IQ, and prior achievement. Another set of variables that sit a little further along the mutability continuum includes factors such as socio-economic status, poverty, parent and child expectations, and parental values. Then there are the more mutable school-related variables such as the curriculum, teacher characteristics, school size, resources and funding, and school climate. Finally there are the most mutable student characteristics such as motivation, beliefs, study habits, health, and truancy. These factors and others like them also vary according to the extent to which educational decision-makers are able to exert influence toward changing them.

The necessarily disjointed nature of existing research means that it is often unclear how the many factors that have been examined either relate to one another or contribute to an overall picture. Multivariate statistical and economic analyses of education outcomes provide some useful insight into these interactions, but often reach inconclusive or contradictory conclusions (e.g., Hanushek 1997, Opdenakker et al. 2002). However, a number of psychologists have proposed models of the learning process that provide some structure to the variables mentioned above (for example, Carroll 1963, Bruner 1966, Cooley and Leinhardt 1975, Bloom 1976, Wiley and Harnischfeger 1974, Harnischfeger and Wiley 1976, Bennett 1978, Gagne 1977, Glaser 1976, Haertel, Walberg, and Weinstein 1983, Reynolds and Walberg 1991, 1992; Walberg 1981; Wang, Haertel, and Walberg 1990, Young, Reynolds, and Walberg 1996, Schreiber 2002, Byrnes and Miller 2007). These models each divide the variable space in different ways, but in general consider a similar range of factors. At the level of the student learning process, they generally consider:

- Cognitive activities, such as memorizing and analyzing
- Affective activities, including motivation, giving effort, and dealing with emotions
- Regulative (or metacognitive) activities, such as planning, monitoring, and reflecting

These learning functions interact with teaching along a continuum from strong to weak teacher control. Thus a teacher can regulate the learning activities completely or share the regulation with students or only loosely regulate. In addition to these student-focused elements, models of the learning process also include a number of factors related to the environment in which the student learns. These include:

- Classroom processes such as time spent on learning, structuring and sequence of learning events, rewards and punishments, instructional methods and quality, and classroom social climate. Many of these are driven by characteristics of the teacher.
- School and district factors, such as classroom time, curriculum, and teaching standards.
- External factors, including home environment, peer environment, and exposure to mass media.
In a few cases, causal links or interactions between variables are suggested in these models, but these relationships are generally underspecified from a computational modeling perspective. The challenge in constructing a conceptual model that is amenable to computational implementation is to 1) select a set of key parameters from these lists that are both a) highly relevant to student outcomes and b) can be specified computationally; and 2) to specify a set of generally realistic interactions among these parameters that can be tested and modified as the modeling process continues.

**Elements of the Conceptual Model**

The conceptual model for this stage of the project was a product of input from education experts (in the form of elicited knowledge and review of the education research literature) combined with insights from the computational modeling team arising from their initial work translating expert input into computational processes. This includes specification of the entities in the model, their characteristics, and their methods of interaction. As the model was implemented, both the conceptual model and the computational model were refined through continuing interactions between the education experts and modeling experts. Although it was not feasible to include all of the elements of the conceptual model in the computational model, these ongoing interactions ensured that the most important elements were included.

The initial conceptual model contained the following entities. Below each entity is a list of attributes associated with it:

- **School system**
  - Schools
  - Students
  - Teachers
- **School**
  - Curriculum (set of classes)
  - Classrooms
  - Individual classes
  - Teachers
  - Students
- **Individual class**
  - Grade level
  - Course name (content)
  - Classroom
  - Teacher
  - Students
  - Size
- **Teacher**
  - Effectiveness (may be derived from data analysis, or may be based on the following sub-attributes)
    - Teaching competence
    - Social competence
    - Subject knowledge
    - Motivation
The student entity is associated with the largest number of attributes, reflecting the fact that tracking student performance is the ultimate goal of the modeling effort. Although there was consensus on the project team that these are the important factors relating to students, the exact function governing the relationship among these attributes has not yet been specified. (This is true of all the entities in the present model.) At this stage, most general student characteristics are treated as immutable attributes, although factors like socioeconomic status and particularly English learner status can change over time. Ability to learn is an aggregate number based on the sub-factors listed, and is also treated as immutable. Finally, prior knowledge (specified in the form of a transcript) plays a more dynamic role in influencing student learning, as it will change over time in response to specific events in the student’s education, such as being assigned to a particular teacher and class during a particular semester.

Teacher effectiveness is the primary way of capturing the influence of teacher-student interactions. A less effective teacher will likely lead to poorer learning and possibly lower grades and test scores for all students taught. However, for example, more highly motivated students may be less influenced by low teacher effectiveness. Due to a lack of data on the sub-attributes contributing to teacher effectiveness, we derive teacher effectiveness directly from San Jose Unified School District data by tracking the average influence of each teacher on the test scores of the students they teach.

The individual class provides the context in which students and teachers interact, and students interact with each other. In each class, student ability to learn and prior knowledge, and teacher effectiveness, are modified by factors such as class size (affecting time available for student-teacher interactions) and aggregate student characteristics. Physical resources available for learning may be included as part of the individual class specification in future model iterations. Note that the “individual class” entity is so named to distinguish classes (a particular gathering of teachers and students) from classrooms (a particular place in a school), which may become important in later iterations of the model. This also helps avoid confusion between educational classes and the technical use of the term “class” to designate any entity in a computational model. The high-level structure of the model, which specifies the baseline interactions between the model entities, is depicted in Figure 2.
Here, the school system and school entities primarily play a coordinating role, assigning students and teachers to classes at the beginning of each learning cycle (which could be a year, a semester, or some other length depending on the school and grade level). In later iterations of the model, school systems and schools may themselves have attributes that directly impact learning outcomes.

The primary output of the computational model is a set of student transcripts, generated through interactions between students, teachers, and classes. These transcripts will make it possible to identify patterns of success and failure that contribute to mathematics education outcomes, which is the ultimate goal of the project. The central organizing process that links the entities in the model is therefore the student transcript process. This process is depicted in Figure 3.
The student transcript process runs once for each class taken by each student in a given learning cycle. The first stage of the process assigns students and teachers to individual classes. This depends on the available classes at the school, teacher qualifications, and (if applicable) classes and grades listed on the student’s prior transcript, which will determine eligibility for certain classes and probability of being assigned to others (for example, whether a student will be assigned to an advanced or remedial mathematics class.) As the model progresses, student choice of classes may be included where applicable.

At the end of each learning cycle, student performance is calculated, based on the student attributes listed above (primarily ability to learn and prior knowledge), teacher effectiveness, and individual class factors (primarily class size and aggregate attributes of students in the class). Aggregate student attributes could be important, for example, if a highly motivated student is assigned to a class with a low average motivation, which could make it more difficult than usual for the highly motivated student to learn, due to peer pressure or classroom dynamics. Performance is calculated on a standardized scale that corresponds to grades in the case of classes, and scores in the case of tests. Tests are assigned to students at various stages of the learning cycle or possibly between cycles; tests scores are calculated similarly to grades.

Finally, after classes and grades are assigned for each student for the current learning cycle, they are added to the student’s cumulative transcript and used as input to future learning cycles.

Organizing the model around this transcript process introduces an important dynamic element. For example, if a given student is assigned to an ineffective teacher for several learning cycles in a row, their
grades and test scores may decline; this in turn may influence the courses they are able to take later on, which will in turn influence their final level of progress through the mathematics curriculum. Similarly, a student with low motivation who is repeatedly placed in classes with more highly motivated students may perform unexpectedly well, which may enable that student to take more advanced classes later on (although they may or may not succeed in those classes depending on their overall learning ability).

**Development and Implementation of the Agent-Based Model**

The conceptual model described above was used to implement an agent-based model of an educational system. Other work on agent-based models in the educational arena includes studies by Tang et al (2006), Harland and Heppenstall (2012) and Araujo and St. Aubyn (2013). Activities associated with a workshop on agent-based modeling and educational systems may be found online (Stanford University, 2011). Morell et al (2010) describe the integration of evaluation and agent-based modeling.

**Specification of the Agent-Based Model**

The conceptual framework can be used to specify primary entities (Actual and potential) for implementation of an ABM. The “Actual” entities are those that are included in the initial implementation, while the “potential” entities listed in smaller font are additional entities that could be added to the simulation model as development of the simulation progresses:

**Student (Ability, Grade Level, age, gender, ethnicity, ESL status, ...)**
Represents a student within the school system

**Teacher (Ability, Gender, Ethnicity, ...)**
Represents a teacher within the school system

**Classroom (Quality)**
Mechanism for grouping one teacher and multiple students together

**School (Curriculum)**
Mechanism for grouping teachers, students, classrooms

**School System (N/A)**
Master “control mechanism” manages simulation and generates reports

Other Objects Within the Modeling Framework include the following:

**Student Transcript**
Record of achievement through the years (e.g. report card)

**Individual Class**
An instance of a specific course, in a specific location, with a specific Teacher, and a specific list of Students
The Process Definition includes three components:

**Input File Generation (Source Data)**
User provides source data that explicitly defines teachers and students
Data files implicitly define schools

**Input File Generation (Define Simulation Starting Point)**
Setup Program reads source data
Inserts definition for curriculum and individual classroom entities
Builds raw input files for the simulation engine

**Execute Model Simulation**
Reads the raw input files created in intermediate step
Simulates progression of students through their K-12 curriculum

---

**Figure 4. Model Representation of School System**
Score Assignment

Score assignment was performed using a weighted average approach. A score function was constructed that represents influence from the following factors:

- The student’s baseline “ability” to learn the material.
- The teacher’s baseline “ability” to teach the material.
- Interactive aspects with classroom (e.g. “average ability” across students).
- How the student performed last year in this class.

The score function is defined as follows:
Class Assignment

Each school in the system has classrooms, teachers, students, and a curriculum. The curriculum is a list of the courses that are to be taught:

- Grade Level (e.g. K-12)
- Course Identifier (integer code for “Math”, “English”, “Social Studies”, ...)
- Course Name (actual string for interpretation “Math”, “English”, ...)
- Number Offered (e.g. “Three different offerings of Second-Grade Math”)

The model expects students to be in classrooms with different teachers and different students as they progress through school. If the students were assigned to the same teacher and classmates every year, this will lead to a degenerate case, and the results will be static and uninteresting. However, it is not uncommon to have some overlap in classmates from year-to-year. Initially, the students were simply assigned using a card shuffling algorithm. In the following illustration, two alternate assignment strategies are considered.

Illustration: Grouping Students by Ability

Given the model specification and the assigned weight function, it is possible to examine different scenarios. As an illustration, we consider two different classroom assignment strategies. Classroom assignment is an area of interest in the educational community. (See, for example, Bosworth and Li, 2013.) In the baseline implementation, students are grouped by ability. In the alternate implementation, students are grouped randomly.

Class assignment is performed by setting up one “Individual Class” for each entry in the defined curriculum. There are multiple “Individual Classes” for each grade/course in the schedule, with the number of individual classes determined by an intermediate setup program in a manner that manages classroom size. Classroom assignment is performed via a simple round-robin “card dealing” method.
Teacher assignment is performed by random assignment. Student assignment is performed using a parameterized assignment that is “partially random”.

We define a parameter $p$ that controls “the amount of randomness” in student assignments. When $p$ is zero, assignments are completely random. When $p$ is one, assignments are designed to group students with similar abilities. We refer to this assignment strategy parameter as the “student-ability grouping bias”.

We now illustrate the mechanics of implementing different grouping strategies by building a table where we construct columns of values to represent the abilities of the students associated with the different strategies.

**Baseline Implementation: Group Students By Ability**

Given $N$ students in a specific grade/school that need to be assigned to $M$ instances of a specific course, we begin by building a table with a column containing student ability values, along with an internal index. The table is sorted in decreasing order by the abilities column (A), as shown in Table 1. The first $N/M$ students are assigned to the first instance, the next $N/M$ students to the next instance, etc. If $N=100$ and $M=4$, this will result in 25 students per class.

**Alternate Implementation: Group Students Randomly**

To group students randomly, we add a second column, (R), filled with random numbers drawn from the interval $[0,1]$. The table is sorted in decreasing order by the (R) column, and students are assigned. It is important to use the same “scale” for these two columns because of the next step.

**Implementation of Student-Ability Grouping Bias**

A third column (K) is added that is filled with the following linear combination of the other columns:

$$K_j = p \cdot A_j + (1 - p) \cdot R_j.$$

The column is sorted, and students are assigned. Consider the following two base cases: Whenever $p = 0$, we get $K_j = R_j$, and whenever $p = 1$, we get $K_j = A_j$. For anything in between, we obtain a linear combination. Note that as $p$ decreases, randomness increases.

<table>
<thead>
<tr>
<th>A</th>
<th>R</th>
<th>K</th>
<th>Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.56</td>
<td>0.34</td>
<td>0.45</td>
<td>...</td>
</tr>
<tr>
<td>0.90</td>
<td>0.02</td>
<td>0.46</td>
<td>...</td>
</tr>
<tr>
<td>0.83</td>
<td>0.86</td>
<td>0.85</td>
<td>...</td>
</tr>
</tbody>
</table>
Table 1. Student Ability Groupings Constructed According to (A) Ability, (R) Random Assignment, and (K) Student-AILITY Grouping Bias

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.79</td>
<td>0.61</td>
<td>0.70</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>0.89</td>
<td>0.49</td>
<td>0.69</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>0.68</td>
<td>0.16</td>
<td>0.42</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>0.54</td>
<td>0.99</td>
<td>0.77</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>0.96</td>
<td>0.67</td>
<td>0.82</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>0.42</td>
<td>0.71</td>
<td>0.57</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
</tbody>
</table>

Experimentation to Demonstrate Grouping Impact

Experimental data sets were derived from the San Jose data as described in further detail below. There were 1306 students, with an estimated ability based on first year language arts scores. There were 819 teachers with an estimated ability based on average before/after math test performance of their students. For the purposes of this experiment, all students and teachers were assigned to a single school, which will allow a more accurate assignment due to the mechanics of the simulation. All students in the simulation enter second grade together, and all teachers are available to teach every class. (Future extensions of the model are anticipated to accommodate more realistic assignments of teachers to classes.) Simulation runs were conducted under two sets of demonstration weights. These weights were chosen to highlight specific aspects of model behavior, and are not intended to reflect realism. Other sets of weights might yield very different results.
The derived data were run through the model using different values of the grouping bias parameter, $p$. For illustration, if we set the weights associated with teacher ability and previous test score to zero, then the resulting student math scores generated by the model will only be a function of student ability and classroom average student ability. Figures 6-10 illustrate sample student paths generated using a specified set of abilities and varying values of $p$. In a similar manner, sample student paths are generated with each of the weights set to one, as shown in Figures 11-15. In this manner, we can examine what happens under different grouping scenarios.

From this demonstration, we make the following observations. Students’ scores are influenced by the average ability of students in each class they take. If students are randomly assigned to classes, all students’ scores tend to cluster around the mean ability of the entire school/district. If students are always grouped with students of similar ability, students’ scores tend to cluster around the mean of that peer group, resulting in a wider distribution of scores overall. In general, high ability students do better, while lower ability students do worse. The initial distribution varies depending on the grouping strategy, but the distribution changes very little in subsequent years. Further work is required to determine whether different weighting schemes may yield different results (in particular, possible trends over time).

This demonstration illustrates how an ABM can be used to explore different scenarios. A wide variety of scenarios can be explored by using different input distributions, different weights, different assignment strategies, and different system structures. As additional variables and complexity are added to the model, there will be extensive capability to explore different scenarios to address different educational questions.

Figure 6. Student Trajectories for Grouping Using Equal Weights for Student Ability and Classroom Average Student Ability (zero weights on Teacher Ability and Previous Test Score) with $p = 1.00$
Figure 7. Student Trajectories for Grouping Using Equal Weights for Student Ability and Classroom Average Student Ability (zero weights on Teacher Ability and Previous Test Score) with $p = 0.75$

Figure 8. Student Trajectories for Grouping Using Equal Weights for Student Ability and Classroom Average Student Ability (zero weights on Teacher Ability and Previous Test Score) with $p = 0.50$
Figure 9. Student Trajectories for Grouping Using Equal Weights for Student Ability and Classroom Average Student Ability (zero weights on Teacher Ability and Previous Test Score) with $p = 0.25$

Figure 10. Student Trajectories for Grouping Using Equal Weights for Student Ability and Classroom Average Student Ability (zero weights on Teacher Ability and Previous Test Score) with $p = 0.00$
Figure 11. Student Trajectories for Grouping Using Equal Weights for Student Ability, Teacher Ability, Classroom Average Student Ability, and Previous Score with p = 1.00

Figure 12. Student Trajectories for Grouping Using Equal Weights for Student Ability, Teacher Ability, Classroom Average Student Ability, and Previous Score with p = 0.75
Figure 13. Student Trajectories for Grouping Using Equal Weights for Student Ability, Teacher Ability, Classroom Average Student Ability, and Previous Score with $p = 0.50$

Figure 14. Student Trajectories for Grouping Using Equal Weights for Student Ability, Teacher Ability, Classroom Average Student Ability, and Previous Score with $p = 0.25$
Figure 15. Student Trajectories for Grouping Using Equal Weights for Student Ability, Teacher Ability, Classroom Average Student Ability, and Previous Score with p = 0.00

Statistical Analysis and Calibration

The Role of Data in the Agent-Based Modeling Approach

Data are incorporated into the modeling process in two ways: 1) Data are used to develop realistic input distributions for the student ability and teacher effectiveness variables. 2) Data are used to calibrate the weights in the model by executing repeated runs of the simulation model and choosing a run that provides the best distributional match of the simulated values to the actual values obtained from the Unified School District Database.

Development of Input and Calibration Data

A demonstration of the use of our prototype model in conjunction with real data was performed using a subset of data from the San Jose Unified School District Database. As would be expected in any school district, there are dynamic effects that occur over time associated with various policy implementations that impact the variables that are collected and the interpretation of the resulting data. Without going into the full complexities encountered, we selected a time frame where test scores were expected to be available and comparable from year to year throughout the time period. This selection was based on discussions with our data expert, Marcy Lauck, who had extensive experience with both the San Jose Unified School District and the contents of the database.
The initial prototype model was implemented as a simple closed cohort system with all students entering at the initial time-point and progressing through the system for a period of seven years. In order to develop distributional information on the initial student ability and the resulting math performance at the end of the progression, we chose to down-select to the population of students who were in the school system for the complete time period from 2004-2005 through 2010-2011 during 2nd through 8th grade. This produced a dataset with initial language arts scores and math scores for each subsequent year for 1306 students. This allowed us to focus on generating realistic behavior without having to address the complexities associated with students entering and exiting the system. This additional complexity could potentially be addressed with future modifications to the ABM and the associated selection of training data. Initial student language scores were used as a proxy to measure initial ability. This provided a measure of intrinsic learning ability that our educational colleagues were comfortable with that was measured separately from the actual mathematics scores that were to be followed over time. Scores were converted to a 0-1 scale by subtracting the minimum score and dividing by the range to generate values suitable for use with the computational model. In a similar manner, student math scores were used to create an output distribution of math performance scores to use as a proxy metric for algebra readiness.

Data on student performance was used as a proxy for measuring teacher ability. Distributional information on a quantity that will be referred to as teacher effectiveness was developed by examining the average change in math scores from one year to the next for all students exposed to a specific teacher in some year. For each teacher, a list was created with all of the students they taught each year, and the differences (positive or negative) in student math scores before and after being taught by that teacher were calculated and rescaled to the same scale used for the scores above. The mean score change associated with that teacher was then used to represent teacher ability. The resulting values were then converted to a 0-1 scale to generate values suitable for use as inputs to the computational model. For this purpose, data included math scores from all students in the system exposed to any teacher who taught math during the specified time period, resulting in a distribution for teacher ability based on a set of 819 teachers.

**Calibration of the Agent-Based Model**

The agent-based model described above includes a number of entities that are specified in the development of the model. A candidate score function was proposed to represent the process by which mathematics performance evolves over time. This function calculates a score that is a weighted function of student ability, teacher ability, and previous score. One can arbitrarily specify values for the weights that will permit execution of the simulation and generation of simulated mathematics scores. However, in order to calibrate the model to the actual data, a calibration approach was developed that involves execution of repeated runs of the model across a space of varying weight parameters. The weights for each run are selected randomly by choosing random values between 0 and 1 for each of the parameters of interest, and then rescaling them to one. This selection process generates vectors of weights representing points that are spread across the multi-dimensional parameter space. Each vector of random weights is used to execute the simulation using the randomly assigned weights for parameters in the score assignment function (student ability, teacher ability, average classroom ability, and previous year’s test score). Two different statistical metrics were proposed: Method 1 computes the correlation between the quantiles of the actual math scores and the quantiles of the model output scores, with higher values indicating closer distributions. Method 2 uses a Kolmogorov-Smirnov test statistic (REFERENCE) to compare distributions of model and actual scores. The KS statistic is based on the
maximum distance between the empirical cumulative distribution functions of the simulated data and the actual data, with smaller KS values corresponding to better matches in distribution. These two metrics can be used to rank the resulting output datasets in order of “closeness” to the distribution of the actual data. The set of weights corresponding to the closest match, using either criterion, can then be used as parameter values in the simulation model. This process of calibration provides estimates of the weight parameters that will result in behavior that best matches the distributional behavior of the actual data, as measured by either correlation or maximum distance between the empirical distribution functions.

These methods were used to compare the results from 100 runs of the ABM with randomly varying weights. Using the correlation method identified the following weights for Student Ability, Teacher Effectiveness, Average Class Ability, and Previous Score associated with the best and worst fits:

Maximum Correlation (best fit): Weights = (.260, .190, .355, .195)
Minimum Correlation (worst fit): Weights = (.015, .376, .436, .173)

Figure 16 shows histograms of the actual data, simulated output for an arbitrary run, simulated data for the run with the highest correlation, and simulated data for the run with the lowest correlation. Note that the histogram for the run with the highest correlation is reasonably similar in shape to the histogram of the actual data, while the bimodal histogram associated with the lowest correlation matches the actual data poorly. Using the KS method identified the following weights associated with the best and worst fits:
Figure 16. Histograms of a) Actual Math Data, b) Arbitrary Run, c) Run Associated with Best Fit, and d) Run Associated with Worst Fit, Using Correlation Metric.
Figure 17 displays histograms of the actual data, simulated output for an arbitrary run, simulated data for the run with the best KS statistic, and simulated data for the run with the worst KS statistic. Note that the histogram for the run with the best KS statistic appears visually to be quite similar to the histogram of the actual data, while the bimodal histogram associated with the worst KS statistic is quite different from the actual data.

![Figure 17](image-url)

Figure 17. Histograms of a) Actual Math Data, b) Arbitrary Run, c) Run Associated with Best Fit, and d) Run Associated with Worst Fit, Using Kolmogorov Smirnov Metric.
Results

An agent-based model has been developed to demonstrate the use of agent-based models for studying educational systems. Model inputs include student ability, teacher ability, classroom, and assignment strategy. The model tracks simulates progress of students over time using math test scores as a measure of algebra readiness and a score function that takes the form of a weighted sum of the inputs and the previous score. There are several technical achievements that were accomplished in the process of exploring the use of agent-based modeling for an educational system. Development of our approach involved several steps including development of a conceptual model, implantation of an agent-based model in a manner that provides for extension and scalability, development of distributional information from actual school district data, development of code to run repeated simulations with varying parameters, development of a calibration approach, and experimentation with different scenarios available within the current structure.

Future Development

While the prototype system described here successfully demonstrates an approach for developing and implementing an agent-based model for an educational system, further development will be required to incorporate more realistic structure and additional variables and associated data into the modeling process to support end-use decision-making. With additional resources, this approach can be extended to accommodate more complex system structures and to incorporate additional input variables, different classroom strategies, etc. to address various use cases and scenarios of interest.

In the proposed calibration methods presented here, the weights were selected by choosing the single best fit from repeated sample runs. Further work is needed to explore the space of weights to assess whether there is actual structure and interpretability across the space of weights, or whether the optimum is simply an empirical solution. Examination of the relationship between different weights could be informative. Experimentation with different weighting schemes could be enlightening and could potentially produce interesting changes or anomalies in the model output.

Clearly, there are a number of potential sources of uncertainty in the ABM modeling process. These include variability in the data used to develop the input datasets, uncertainty arising from the inherent stochastic nature of human behavior, uncertainty in the specification of the model structure and score function, as well as variability present in the math test scores used to perform calibration. These uncertainties will then be propagated to the resulting simulation model outputs and weight estimates. Assessment of these different types of uncertainty warrants further attention in future model development efforts. Previous work on uncertainty and error for agent-based models may be found in Evans (2012).

A number of different use cases can be envisioned along with corresponding development of the model. The assessment of the impact of different types of grouping strategies, as illustrated earlier, could provide a useful thought experiment/discussion generator for administrators. Education researchers could use such results to assess model assumptions and suggest alternatives. A similar strategy could be
used to assess the impact of efforts to increase teacher effectiveness. Tracking the paths of individual students over time as the model progresses may be useful for exploring potential outcomes and how student progress in math is influenced by teachers, classmates, and other factors. If factors can be identified that promote movement from lower relative performance to higher relative performance, such factors could potentially be used to achieve distributions with overall improved performance.

**Summary and Conclusions**

This project has developed a conceptual framework for modeling of an education system and demonstrated the use of agent-based modeling techniques to implement the structures specified in the framework. The demonstration involves implementation of an ABM informed by statistical data. Data analysis procedures were developed to produce realistic sample input distributions. A statistical calibration approach was developed to select values for the simulation model parameters to produce distribitional behavior that is consistent with the actual data. While this type of modeling approach is relatively complicated, it can be particularly useful in situations where there is a desire to capture interactive behavior between model components in large, heterogeneous systems. (See O’Sullivan et al, 2012).

**Acknowledgements**

The authors wish to thank the following individuals for their help and encouragement in this modeling effort: Gordon Freedman from the National Laboratory for Education Transformation (NLET) for facilitating the Unified School District Database access with Follett Software Company, Marcy Lauck from NLET and formerly the San Jose Unified School District, for assistance with data acquisition, extraction, and interpretation, Bill Rust from Los Alamos National Laboratory for programming assistance for the varying weight runs, Paul Resta, Ken Tothero, and Tom Benton from the University of Texas at Austin for complementary visualization efforts that provided useful insights, John Gargani, Gargagni + Company, for helpful comments and suggestions, and Caroline Hardeman from NLET for her enthusiastic input and support throughout this project.

**References**


Stanford University (2011), Agent-Based Modeling of Educational and Social Systems, 
Tang, Y., Parsons, S., and Sklar, E. (2006), Modeling human education data: from equation-based 
modeling to agent-based modeling. MABS’06 Proceedings of the 2006 International Conference
(Eds.), Psychology and education (pp. 81-110). Chicago: National Society for the Study of
Education.
instruction, major educational vehicles. Educational Researcher, 3, 7-12.
## Appendix 1. Elicited Factors Contributing to Learning Outcomes

<table>
<thead>
<tr>
<th>Factor</th>
<th>Agent(s)</th>
<th>Definition</th>
<th>Indicators</th>
<th>Data?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Opportunity to learn</strong></td>
<td>School, teacher</td>
<td>Time engaged in learning activities</td>
<td>Doing/not doing homework</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Time in school day</td>
<td>Not in database, but available</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Classes, attendance/truancy</td>
<td>In database?</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Classroom management and teacher experience (efficient use of time)</td>
<td>No, but may be tied to teacher experience</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Policy/curriculum mandates (time spent on specific subjects)</td>
<td>Not in database, but available</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Teacher attendance?</td>
<td>No (union rules)</td>
</tr>
<tr>
<td><strong>Instructional content, structure, and timing</strong></td>
<td>School, district, state</td>
<td></td>
<td>Course catalogs</td>
<td>Available, but not in database</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>State blueprints</td>
<td>Available, but not in database</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Pacing guides</td>
<td>Available, but not in database</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Variations between teachers</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Curriculum documents (similar to state blueprints?)</td>
<td>Available, not in database, may vary locally</td>
</tr>
<tr>
<td><strong>Teacher ability/ instructional quality</strong></td>
<td>Teacher</td>
<td>Standards (guidelines that specify best teaching practices in general)</td>
<td>Available, but not in database. Changed over time span represented in database.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Whether teachers follow standards</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Performance of past students (stickers, sliders, gainers – i.e. how many students progressed on pace, how many slid back, and how many gained). Can be viewed as “value added” by teacher</td>
<td>Can calculate from test scores in database.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Teacher preparation – degrees, certificates, etc.</td>
<td>May be in database</td>
<td></td>
</tr>
<tr>
<td>School resources</td>
<td>School</td>
<td>Material, personnel, finances, etc.</td>
<td>Teacher personal characteristics – personality, communication skills</td>
<td>No (classroom assessments of teachers are not in database – again, union rules)</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Years of experience (data shows effectiveness climbs steeply in first few years, then levels off or starts to drop a little)</td>
<td>Probably in database</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Many other characteristics; studies have shown many of the indicators listed here are actually not highly correlated with teaching quality. See Michael Strong’s book on this subject.</td>
<td>Possibly</td>
</tr>
<tr>
<td>School size might have an effect</td>
<td>School</td>
<td>Material, personnel, finances, etc.</td>
<td>School size might have an effect</td>
<td>Available, maybe not in database?</td>
</tr>
<tr>
<td>Dollars per student (correlation with student success is unclear)</td>
<td>School</td>
<td>Material, personnel, finances, etc.</td>
<td>Dollars per student (correlation with student success is unclear)</td>
<td>Available, but not in database</td>
</tr>
<tr>
<td>Class size (correlation unclear)</td>
<td>School</td>
<td>Material, personnel, finances, etc.</td>
<td>Class size (correlation unclear)</td>
<td>In database</td>
</tr>
<tr>
<td>Availability of school activities</td>
<td>School</td>
<td>Material, personnel, finances, etc.</td>
<td>Availability of school activities</td>
<td>Student self-report of activity participation available in survey results – Marcy can include in future</td>
</tr>
<tr>
<td>Physical infrastructure/school condition (impact has not been studied)</td>
<td>School</td>
<td>Material, personnel, finances, etc.</td>
<td>Physical infrastructure/school condition (impact has not been studied)</td>
<td>No</td>
</tr>
<tr>
<td>Socioeconomic status</td>
<td>Student</td>
<td>Demographic characteristics that could impact school success</td>
<td>Socioeconomic status</td>
<td>Free/reduced lunch participation (in database)</td>
</tr>
<tr>
<td>Gender</td>
<td>Student</td>
<td>Demographic characteristics that could impact school success</td>
<td>Gender</td>
<td>In database</td>
</tr>
<tr>
<td>Race</td>
<td>Student</td>
<td>Demographic characteristics that could impact school success</td>
<td>Race</td>
<td>In database</td>
</tr>
<tr>
<td>Neighborhood (unclear significance)</td>
<td>Student</td>
<td>Demographic characteristics that could impact school success</td>
<td>Neighborhood (unclear significance)</td>
<td>Student zip codes – in database</td>
</tr>
<tr>
<td>Parent education level (stands in for a number of possible indicators with no data, including # hours children out of community, books and reading materials in home, parent engagement, arrive ready for school, learning outside of school, parents value education.)</td>
<td>Student</td>
<td>Demographic characteristics that could impact school success</td>
<td>Parent education level (stands in for a number of possible indicators with no data, including # hours children out of community, books and reading materials in home, parent engagement, arrive ready for school, learning outside of school, parents value education.)</td>
<td>Parent education level is self-reported in survey results – Marcy can include in future</td>
</tr>
<tr>
<td>Characteristics of other students from same neighborhood (school achievement, attendance, truancy, socioeconomic status, etc.)</td>
<td>Student</td>
<td>Influences of fellow students/friends</td>
<td>Characteristics of other students from same neighborhood (school achievement, attendance, truancy, socioeconomic status, etc.)</td>
<td>Other students from same zip code and their demographics and academic records (in database)</td>
</tr>
<tr>
<td>Characteristics of other students in same classes (see above)</td>
<td>Student</td>
<td>Influences of fellow students/friends</td>
<td>Characteristics of other students in same classes (see above)</td>
<td>Other students in same classes and their demographics and academic records (in database)</td>
</tr>
<tr>
<td>Category</td>
<td>Type</td>
<td>Description</td>
<td>Source</td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------</td>
<td>------------</td>
<td>------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Student motivation/perseverance</td>
<td>Student</td>
<td>Willingness to commit time to school work</td>
<td>Survey results, Marcy can include</td>
<td></td>
</tr>
<tr>
<td>Belief in need to work hard</td>
<td></td>
<td></td>
<td>Survey results, Marcy can include</td>
<td></td>
</tr>
<tr>
<td>Future orientation</td>
<td></td>
<td></td>
<td>Predefined index of several factors reported in survey results, Marcy can include</td>
<td></td>
</tr>
<tr>
<td>Developmental assets/intrinsic and extrinsic (search on this)</td>
<td></td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRESS index</td>
<td></td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IQ</td>
<td></td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student aptitude/ability</td>
<td>Student</td>
<td>Student ability to learn in a given environment (partly innate, partly learned)</td>
<td>In database</td>
<td></td>
</tr>
<tr>
<td>GATE (gifted and talented) designation – based on testing, related to IQ.</td>
<td></td>
<td>In database</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Previous performance (grades, tests)</td>
<td></td>
<td>In database</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Special needs designation</td>
<td></td>
<td>In database</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prior knowledge and skills</td>
<td>Student</td>
<td>Skills and knowledge that facilitate learning</td>
<td>In database – test scores</td>
<td></td>
</tr>
<tr>
<td>English as a second language designation (ESL) and reclassification dates</td>
<td></td>
<td>In database</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home language</td>
<td></td>
<td>May be in survey results (Marcy)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language proficiency (influences learning in many subjects, including STEM)</td>
<td></td>
<td>In database – test scores</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other measures of proficiency, e.g. in math?</td>
<td></td>
<td>In database – test scores</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitive difficulty</td>
<td>Class</td>
<td>Inherent difficulty of course material (slows all students down proportional to their abilities)</td>
<td>May be in database (course names or separate designator?) or can be found in course catalogs.</td>
<td></td>
</tr>
<tr>
<td>Courses designated AP, honors, a-g. (A-g is a University of California standard for college prep)</td>
<td></td>
<td>May be in database (course names or separate designator?) or can be found in course catalogs.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whether in “zone of proximal development” for student (best learning occurs just beyond current abilities)</td>
<td></td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prerequisite knowledge/skills</td>
<td>Class</td>
<td>Prior knowledge/skills that enhance ability to learn course material</td>
<td>Possibly in course catalogs or other documents, if not in database.</td>
<td></td>
</tr>
<tr>
<td>Course prerequisites (these may determine classes available to student rather than their performance in classes)</td>
<td></td>
<td>Possibly in course catalogs or other documents, if not in database.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>General knowledge/skills related to subject matter</td>
<td></td>
<td>No</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 2. Internal Literature Review of Models and Variables Related to Student Math Outcomes

Introduction

The benefit of using an agent-based modeling approach to assess the predictors of student achievement in mathematics is that it allows us to examine a large set of factors in a single study. In most previous work researchers have examined individual or small sets of variables. The purpose of this paper is to identify variables that have been demonstrated to have a relationship with mathematics achievement, so that we may consider how best to use that information to construct a conceptual framework that will inform our study of SJUSD data.

The necessarily disjointed nature of existing research means that we do not have a clear idea of how the many factors that have been examined either relate to one another or contribute to an overall picture. Furthermore, we are not able to discern the relative importance of the various factors that appear to be predictive of student achievement.

So what are the variables that have been the focus of previous research and how do they fit together and relate to one another? A review of studies from the past 50 years or so reveals a wide variety of factors that have demonstrated some connection to student achievement in general or to mathematics achievement in particular. Since our goal is to provide practical information to school, district, and policy personnel in order to help transform education, one possible way of organizing the variables from these studies is according to their susceptibility to change (their mutability). A second related consideration might be the degree of control a school administrator or policy maker may have to change these variables. Of course, not all these variables will exist in our database.

A number of variables describe intrinsic student characteristics at the low end of the mutability scale. Most prominent of these are gender and race, to which we may wish to add aptitude, IQ, and prior achievement. Another set of variables that sit a little further along the mutability continuum includes factors such as socio-economic status, poverty, parent and child expectations, and parental values. Then there are the more mutable school-related variables such as the curriculum, teacher characteristics, school size, resources and funding, and school climate. Finally there are the more mutable student characteristics such as motivation, beliefs, study habits, health, and truancy. These factors and others like them also vary according to the extent to which educational decision-makers are able to exert influence toward changing them.

Different researchers, depending on their various perspectives, have presented holistic conceptions of classroom learning using different configurations of these and similar variables. We can begin by looking at the main psychological models of school learning that have been developed over the years (since they are the most numerous), followed by a glimpse of economics and statistical models. This is by no means an exhaustive examination.

Psychological Models
John Carroll (1963), known for his study of foreign language aptitude, proposed a general model of school learning based on the assumption that learning is contingent upon the time students are willing and able to invest in the learning process. His model has five main constructs:
a) *Aptitude* – the amount of learning time necessary for a student to master an objective under optimal learning conditions (more aptitude = less time)
b) *Perseverance* – the amount of time a student is willing to invest in mastering the objective
c) *Ability to comprehend instruction* – related to general or verbal IQ
d) *Opportunity to learn* – the amount of time a teacher allots for learning the content
e) *Quality of instruction* – the organization of instruction for ease of acquisition by students

We can see that the first three constructs can be measured by amounts of time, while the last two require analysis of instruction.

Bruner (1966) outlined a normative theory of instruction organized around four requirements:

a) *Implanting a predisposition toward learning* – motivation for learning
b) *Structuring the body of knowledge to be taught* – differs according to learners, previous knowledge, and nature of subject
c) *Sequencing the presentation of materials to be learned* – ideally from enactive to iconic to symbolic
d) *Specifying the nature and spacing of rewards and punishments* – includes reinforcement and intrinsic and extrinsic rewards as well as feedback

Cooley and Leinhardt (1975) developed a process model that centers on the relationship between school practices and school performance. They looked at outcomes related to academic achievement and attitudes toward school, peers, and teachers. These were:

a) *Initial abilities* – general ability, prior achievement, attitudes towards school, peers, and teachers
b) *Opportunity* – the amount of time students could work on specific content
c) *Motivators* – the student behaviors and attitudes that promote learning activities (included both internal, such as choice of leisure time activities, or external, such as teacher praise)
d) *Structure* – organization and sequencing of curriculum, specificity of objectives, matching of students and curriculum
e) *Instructional events* – the instructional interactions of an interpersonal nature: content, frequency, quality and length

Bloom (1976) developed a model influenced strongly by Carroll’s. It features two types of learning prerequisites and quality of instruction:

a) *Learner’s cognitive entry behaviors* – similar to Carroll’s aptitudes.
b) *Affective entry characteristics* – include attitude towards subject matter and school, and self-concept as a learner
c) *Quality of instruction* – Cues (clarity of presentation and explanation of learning activities), reinforcements (praise and blame, encouragement etc.), feedback and correctives, and participation (time on task)

Bloom’s model has three learning outcomes, achievement, affective behaviors, and improved rate of learning.

The Harnischfeger-Wiley model of school learning consists of a framework with six components divided among three categories (Wiley & Harnischfeger, 1974; Harnischfeger & Wiley, 1976) and has some basis in both Carroll’s and Bloom’s work:
productivity factors can be divided into three sets.

a) Student aptitude – student ability and/or prior knowledge, the developmental stage of a student, and a student’s motivation.

b) Instruction – quantity of instruction (amount of time), and quality of instruction.

c) Environment – class climate, home environment, peer environment, and exposure to mass media.
The underlying rationale of the model is that psychological attributes and their proximate environments influence cognitive, behavioral, and attitudinal outcomes of education.

More recently Schreiber (2002) used the H-W model as a basis for a study of TIMMS data, and developed a model of his own focusing on institutional and student factors. Schreiber’s construct used the Harnischfeger-Wiley factor of background for the constructs of Institutional Factor (manifested by resources, aggregate parent education, incidents, and school size) and the Student Factor (manifested by highest parent education, gender, math/physics student, attitude toward mathematics, natural talent belief, and hard work belief). The Teaching-Learning Process factor produced the In-class pursuits (manifested by active responding composite and passive responding composite) and Out-of-Class Pursuits (manifested by television watching, employment, sports, and studying mathematics). Finally Acquisition produced the construct Achievement (manifested by an advanced mathematics score).

A recent attempt at an integrative theory of academic achievement in mathematics and science is the opportunity-propensity model of Byrnes and Miller (2007). They distinguish three categories of factor:

a) Distal factors – Factors that enable or explain the emergence of opportunities to learn or propensities. Includes SES, parental expectations, children’s expectations, and prior educational experience.

b) Opportunity factors – Culturally defined contexts in which an individual is presented with content to learn (e.g., by a teacher or parent, an author, a narrator of an educational TV program, etc.) or given opportunities to practice skills.

c) Propensity factors – any factors that relate to the ability or willingness to learn content once it has been exposed or presented in particular contexts. Includes intelligence, aptitude, cognitive level, and pre-existing skills, along with motivation, interest, self-efficacy, values and competence perceptions.

Economics Models

The economist’s typical approach is to focus on school resources and their relationship to student achievement. The underlying principle is that the greater and better the resources the greater and better will be the student outcomes. Hanushek (e.g., 1997) is well known for his focus on and reviews of these kinds of studies. Hanushek reviewed over 400 studies starting with the well-known “Coleman Report” (Coleman et al., 1966). Hanushek identifies at least three distinct categories of resources among these studies, of which the first receives the most attention:

a) The real resources of the classroom (teacher education, teacher experience, and teacher-pupil ratios

b) Financial aggregates of resources (expenditure per student and teacher salary)

c) Measures of other resources in schools (specific teacher characteristics, administrative inputs, and facilities)

Table 1 is taken from the Hanushek paper and presents a summary of the results of the studies he reviewed. As we can see, under the real resources category, only 9% of the studies looking at teacher education and 15% of those looking at teacher-pupil ratios find significant positive effects on student performance. Balancing these somewhat low numbers are the statistically significant negative findings (5% and 13% respectively). The studies with statistically insignificant findings that report directionality show a fairly even split between positive and negative.
The percentages for teacher experience are higher, but a full 71% indicate worsening performance with experience or less confidence in a positive effect. The conclusion from the numbers displayed in Table 1 is that adding any of these kinds of resources to schools will probably not lead to any increases in student achievement. The same is true for the financial aggregates. Hanushek further analyzes the data from the studies he reviewed and his resulting simple interpretation is that there is no strong or consistent relationship between school resources and student performance.

Another way of interpreting Hanushek’s findings is that a policy of increasing resources would, on average, not lead to an increase in student achievement. Given the right incentives, local school administrators may be able to use additional resources in ways that, on average, improve student achievement, but only if the proper incentives are in place. On the other hand, decreasing resources below some unknown limit would hurt student achievement.

**Statistical Models**

Opdenakker et al. (2002), like many others have done, opted for possibly the simplest structure in their multi-level statistical model for examining the effect of schools and classes on mathematics achievement in Belgium. It is not based on any conceptual model or theory. They separated:

a) **student-level** variables (initial cognitive ability, prior mathematics achievement, family SES, achievement motivation, immunity to stress, sex, and language spoken at home)

b) **class-level** variables which were essentially the aggregates of the student level variables (mean initial cognitive ability, mean SES, mean achievement motivation, mean immunity to stress, proportion of girls, and proportion of students who spoke Dutch at home)

c) **school-level**, which, in addition to the aggregates of the class level variables, included several variables related to educational process, counseling, and school leadership.

These were: attention towards differences between students, focus on discipline and subject matter acquisition, orderly learning environment, use of test results to improve teaching, extent to which formal structure and regulations were typical of the school, evaluation of the functioning of the school by teachers and by students, attention of the school principal to pedagogical aspects, student coaching at school, average level of structured teaching, average of attention to high- and low-achieving students, average level of consultation between teachers, average level of feedback on study results, average proportion of intellectually challenging questions in a usual class test of

---

**Table 1**

**Percentage Distribution of Estimated Effect of Key Resources on Student Performance, Based on 377 Studies**

(From Hanushek, 1997)

<table>
<thead>
<tr>
<th>Resources</th>
<th>Number of estimates</th>
<th>Statistically significant</th>
<th>Statistically insignificant</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>Real classroom resources</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher-pupil ratio</td>
<td>277</td>
<td>15%</td>
<td>13%</td>
</tr>
<tr>
<td>Teacher education</td>
<td>171</td>
<td>9%</td>
<td>5%</td>
</tr>
<tr>
<td>Teacher experience</td>
<td>207</td>
<td>29%</td>
<td>5%</td>
</tr>
<tr>
<td>Financial aggregates</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher salary</td>
<td>119</td>
<td>20%</td>
<td>7%</td>
</tr>
<tr>
<td>Expenditure per pupil</td>
<td>163</td>
<td>27%</td>
<td>7%</td>
</tr>
</tbody>
</table>

*Note: Source - Hanushek’s tabulations.*
mathematics, average level of opportunity to learn mathematics, and attention to individual students by mathematics teachers. Schools were also categorized as public or catholic, and by the study programs they offered.

This brief overview of some of the existing models that have been used to describe the relationship between potentially predictive factors and student achievement may serve as a basis for a discussion about the conceptual approach we take to our own study. Most of the models reviewed began with some kind of logical or theoretical perspective on learning and then proceeded to collect and/or analyze data to establish which factors actually show empirical evidence of a connection with student outcomes. In our case we might proceed by discussing whether we wish to align ourselves with any particular theoretical model of learning, or take a bottom up approach and perhaps develop our own theory, or apply a combination of the two.