

There Once Was a 9-Block ...- A Middle-School Design for Probability and Statistics

Dor Abrahamson
University of California, Berkeley

Ruth M. Janusz
Nichols Middle School, Evanston, IL

Uri Wilensky
Northwestern University

Journal of Statistics Education Volume 14, Number 1 (2006),
www.amstat.org/publications/jse/v14n1/abrahamson.html

Copyright © 2006 by Dor Abrahamson, Ruth M. Janusz and Uri Wilensky, all rights reserved. This text may be freely shared among individuals, but it may not be republished in any medium without express written consent from the authors and advance notification of the editor.

Key Words: Computers; Education; Mathematics; Sample; Statistics.

Abstract

ProbLab is a probability-and-statistics unit developed at the Center for Connected Learning and Computer-Based Modeling, Northwestern University. Students analyze the combinatorial space of the 9-block, a 3-by-3 grid of squares, in which each square can be either green or blue. All 512 possible 9-blocks are constructed and assembled in a “bar chart” poster according to the number of green squares in each, resulting in a narrow and very tall display. This combinations tower is the same shape as the normal distribution received when 9-blocks are generated randomly in computer-based simulated probability experiments. The resemblance between the display and the distribution is key to student insight into relations between theoretical and empirical probability and between determinism and randomness. The 9-block also functions as a sampling format in a computer-based statistics activity, where students sample from a “population” of squares and then input and pool their guesses as to the greenness of the population. We report on an implementation of the design in two Grade 6 classrooms, focusing on student inventions and learning as well as emergent classroom socio-mathematical behaviors in the combinations-tower activity. We propose an application of the 9-block framework that affords insight into the Central Limit Theorem in science.

1. Introduction

The National Council of Teachers of Mathematics has set the standard that middle-school students should be able to “compute probabilities for simple compound events, using such methods as organized lists, tree diagrams, and area models” ([NCTM 2004](#)). We ask ourselves which activities may foster opportunities for students to learn the methodology of combinatorial analysis through partaking in the development of combinatorial-analysis methods. Furthermore, we are attempting to create activities that help students contextualize combinatorial analysis within the broader conceptual framework of probability and statistics, as a means of grounding a comprehensive, flexible, and connected understanding of these related topics.

This paper overviews our design work and reports on an implementation of this design in a middle-school classroom, focusing on the combinatorial-analysis activity. Following an overview of the design, we introduce the types of mathematical problems that are achievement goals for our students, and then we use these problems to explain the principles, structure, and activities of the combinations tower and some of the computer models. The article continues with a description of the lessons in Ms. Janusz’s two 6-grade classrooms ($n = 21$, $n = 19$), interleaved with our reflection on what we believe students learned in these activities. Next, we propose how combinatorial analysis might help students understand statistical distributions in populations, for instance why it is that the heights of all 6-grade girls in the U.S.A. are distributed in the form of a bell-shaped or ‘normal’ curve. The paper ends with an outline of our future research.

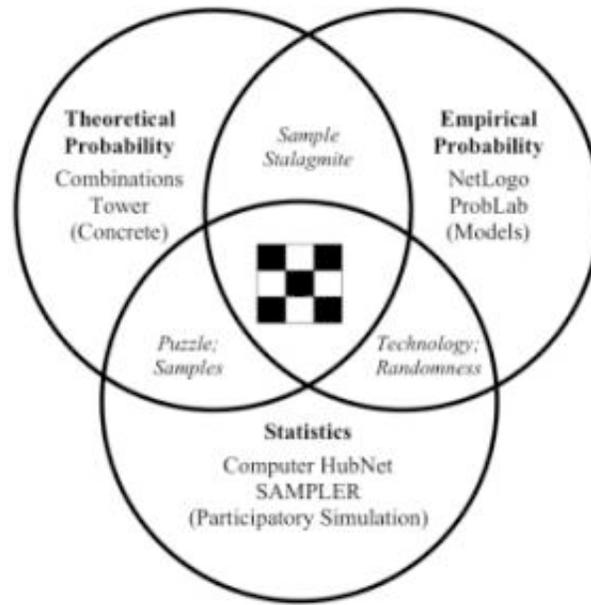
Readers interested in the design-research framework that informed this study are referred to [Abrahamson and Wilensky \(2004c\)](#).

1.1 Overview of the Design

For a 6-grade unit on probability and statistics, we designed the combinations tower activity on combinatorial analysis. This hands-on learning-by-constructing activity lasts approximately three double periods (about 80 minutes each). The learning objective of the activity is for students to build experiential meaning for probability and statistics. Specifically, we wanted students to connect theoretical and empirical probability and to prepare them for a miniunit in statistics (see Figure 1, for our domain analysis in designing the whole unit). These mathematical ideas are usually taught in later grades, yet we were interested in investigating younger students’ preparedness for learning these ideas. The activity is part of the ProbLab curricular unit ([Abrahamson and Wilensky 2002, 2004a, b, c](#)) which in turn is part of the Connected Probability project ([Wilensky 1993, 1995, 1997](#)).

In order to help students compare, contrast, and connect the three identified conceptual pillars of the domain —theoretical probability, empirical probability, and statistics— we designed a new mathematical object as a *bridging tool* ([Abrahamson 2004, Fuson and Abrahamson 2005](#)) between these pillars. This object is the “9-block,” a 3-by-3 grid in which each square can be either green or blue (see [Figure 1](#), center, for an example of a 9-block combination). As we now explain, the 9-block features in different media and different activities as the “math-thematic” bridging tool – in its different guises, the 9-block functions either as a template for conducting combinatorial analysis ([Figure 2a](#)), a stochastic device generating random outcomes ([Figure 2b](#)), or a sampling tool ([Figure 2c](#)). These multiple roles of the 9-block are expressed through the following classroom activities. The combinations tower (see [Figure 2a](#)) is a construction of all the 512 different 9-block combinations determined through combinatorial analysis, a mathematical methodology often applied in the solution of problems in theoretical probability. For the empirical probability part of our

unit, we prepared interactive computer models that generated random 9-blocks (see for example [Figure 2b](#)). For the statistics part of our unit, we developed a computer-based learning environment in which 9-blocks were samples taken out of a giant population grid (see [Figure 2c](#), for the computer interface of an individual student who is taking 9-block samples from this population). Our design is for teachers to help students untangle and understand the conceptual pillars of the domain – theoretical probability, empirical probability, and statistics – by using the 9-block as a basis for cross-referencing, comparing, and contrasting these pillars.

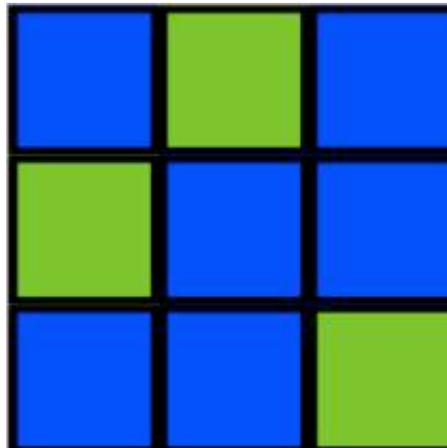


[Figure 1](#)

Figure 1. This diagram illustrates the overall plan of ProbLab, a curricular unit in probability and statistics. The “9-block,” in the center of the figure, is the math-thematic object of this unit that helps students understand and connect theoretical and empirical probability and statistics. This paper focuses on the top-left space of this diagrams - the Combinations Tower that is a combinatorial-analysis classroom project - and discusses its connections to the unit.



[Figure 2a](#)



[Figure 2b](#)



[Figure 2c](#)

Figure 2. The 9-block ties together our designs for: (a) combinatorial analysis and theoretical probability (on left); (b) empirical probability (in the center); and (c) statistics (on the right). On the left, Ms. Janusz and her 6th-grade students stand by the combinations tower constructed from paper-and-crayon 9-blocks that they created and cut out of blank-grid worksheets. In the middle is a computer generated random 9-block. On the right are 9-block samples from a computer-based population of thousands of squares that are each either green or blue.

1.2 The 9-Block as a Bridging Tool

Consider the following problems. Problem 1a: “How many different combinations are there for a 3-by-3 array of squares, if each square can be either green or blue?” Problem 1b: “If all these combinations were put in a hat and one of them were drawn out, what is the chance of getting a combination with exactly 4 green squares?” These types of problems, which are canonical in reform-based mathematics curricula for late-elementary and middle school, represent two complementary aspects of studying probability. The first problem is in combinatorial analysis (combinatorics), a branch of discrete mathematics that studies the space of all possible arrangements of a set of variables that can each take on some specified values. The second problem is in probability, another branch of discrete mathematics, that investigates the likelihood and distribution of events. The two problems are closely related: once we determine how many different combinations there are in total and how many out of all these combinations have exactly 4 green squares, the ratio between the subgroup and the total is the probability of drawing any four-green combination from the hat.

There are 512 different combinations of 3-by-3 arrays (“9-blocks”) made up either of green or blue squares (2^9). Of these 512 combinations, there are 126 different combinations with exactly 4 green squares, as derived from the binomial formula, so there is about one quarter of a chance ($126 / 512 = \sim 0.25$) so as to draw a “four-green 9-block” out of a hat that contains all 512 different combinations. Thus, the combinatorial analysis creates a sample space of all events that each have the same likelihood of occurring, and by identifying and counting up a subgroup (class of events) within that space we can anticipate the results of a probability experiment.

Whereas combinatorial analysis and computing probabilities are middle-school learning goals, the mathematical formulas and procedures for calculating these combinations and likelihoods may be difficult for many middle-school students (see the scores of U.S.A. Grade 12 students, [NCES 2004](#)). For example, only 3% of 12th-grade students solved correctly the following problem ([NCES 2004](#)): “A fair coin is to be tossed three times. What is the probability that 2 heads and 1 tail in any order will come up?” In designing the combinations tower that is the focus of this paper, our rationale was that students may need more support in building sense for and connecting between combinatorial analysis and probability experiments (see also [Wilensky \(1997\)](#)). Using paper, crayons, and glue, students collaboratively construct a classroom visual representation of the combinatorial sample space of all 9-blocks and in doing so they “construct” their own meaning for probability (see [Papert \(1991\)](#), for a pedagogy that relates these two kinds of constructing; see also [Eizenberg and Zaslavsky \(2003\)](#), on the advantage of collaboration for learning combinatorics).

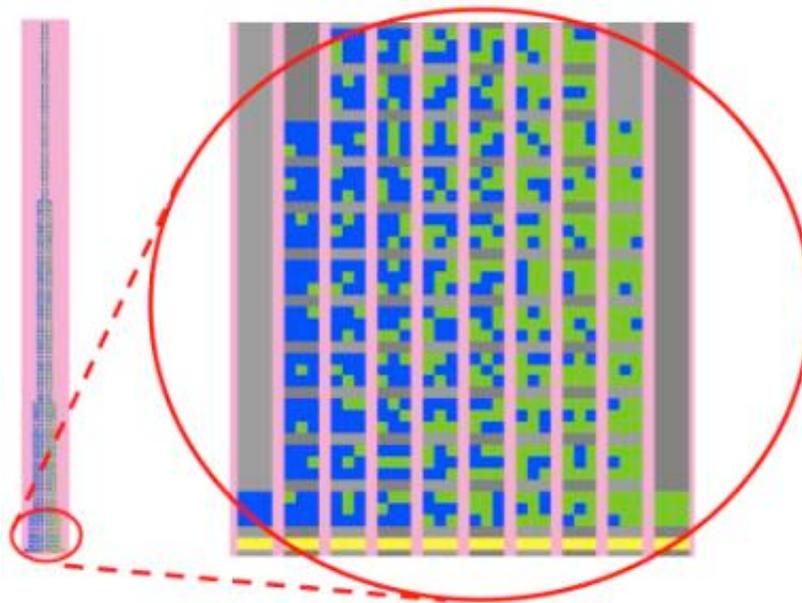
Investigating the 9-block offers opportunities for engaging in combinatorial analysis of a compound event. Within the context of probability experiments, any random 9-block outcome is a compound event, because its identity – its particular green/blue pattern – is a specific outcome configuration of its nine independent

squares. For example, a 9-block with only one green square in the top-left corner is the compound event configured of nine concurrent independent outcomes: “top-left square is green AND top-middle square is blue AND top-right square is blue, ..., AND bottom-right square is blue.” Note that unlike 9 tossed coins, which would land “all over the place,” the inherent spatiality of the 9-block fixates the location of the independent outcomes and thus creates unique visual identities (patterns) for each of the 512 possible combinations. (A plausible compound-event probability problem might be the following: “A fair coin is tossed 9 times; What is the probability that 4 heads and 5 tails in any order will come up?” An analogous problem would be, “9 fair coins are tossed at the same time. What is the probability that 4 heads and 5 tails in any order will come up?” Understanding the mathematical analogousness of these two situations – the sequential and the simultaneous – is in and of itself non trivial, yet we will focus on the latter situation, where outcomes occur concurrently.) Attending to these unique identities of 9-blocks, such as in the context of constructing their combinatorial space, may encourage naming and classifying these combinations, e.g., as in terms of k -subsets. For instance, a given 9-block, which may be interpreted as bearing a specific identity, e.g., “the 9-block with a row of three green squares on top,” may also be interpreted as “one of the possible patterns for exactly three greens,” i.e., as a member of the k -subset “all the three-green 9-blocks.” Appreciating the difference between a unique element in the combinatorial space and a class of elements is conducive to understanding that whereas each element is equally likely to occur, classes of elements differ in frequency (compare, for example the $1/512$ frequency of the zero-green class to the $9/512$ frequency of the one-green class). In contrast, in the case of 9 tossed coins one cannot readily attend to the individual identities of each coin, and so it is more difficult to “discover” methods of combinatorial analysis. In sum, the 9-block form is designed to scaffold student insight into the systematicity and rigor of combinatorial analysis by serving as a template that helps students initially see, create, and organize the combinatorial space.

We believe that there are educational benefits to the combinations tower activity even if computers are not incorporated into the unit. At the same time, the computer-based components can inform students’ strategies because computer-based probability experiments foreground the difference between rigorous combinatorial analysis and random sampling. Whereas the focus of this article is the combinations tower activity, we discuss the computer work so as to demonstrate some advantages of a mixed-media learning environment ([Abrahamson, Blikstein, Lamberty, and Wilensky, 2005](#)).

1.3 The Combinations Tower as “The Shape of Things to Come”

The combinations tower is a “combinatorial sample space” (see [Figure 3](#)). It is a hybrid histogram that shows not only the heights of each column as in regular histograms—it stacks the collection of the combinations themselves as elements in their respective columns (as in a Galton Box). In these columns, the combinations are grouped according to how many green squares are in each combination. A visual comparison between the heights of columns in this histogram may support a sense of the columns’ relative proportion in the sample space and therefore of expected relative frequencies in probability experiments. For instance, the one-green column has 9 permutations and the two-green column has 36 permutations, and this means that a two-green combination is 4 times as likely as a one-green combination. The underlying powerful idea is that combinatorial analysis can help us anticipate the results of an empirical-probability experiment. ProbLab is designed to foster this powerful idea linking theoretical and empirical probability. In addition, the unit links probability and statistics: the combinations tower represents the chances of sampling a 9-block combination with exactly 0, 1, 2, ..., 8, or 9 green squares, respectively, from a matrix “population” with thousands of randomly-distributed squares, where 0.5 of the squares are green and .5 are blue (see [Figure 2](#) and see [section 1.4](#), below).

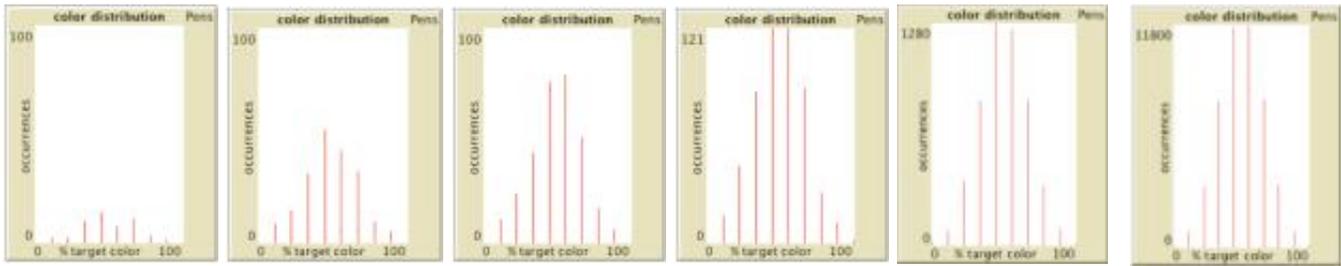


[Figure 3](#)

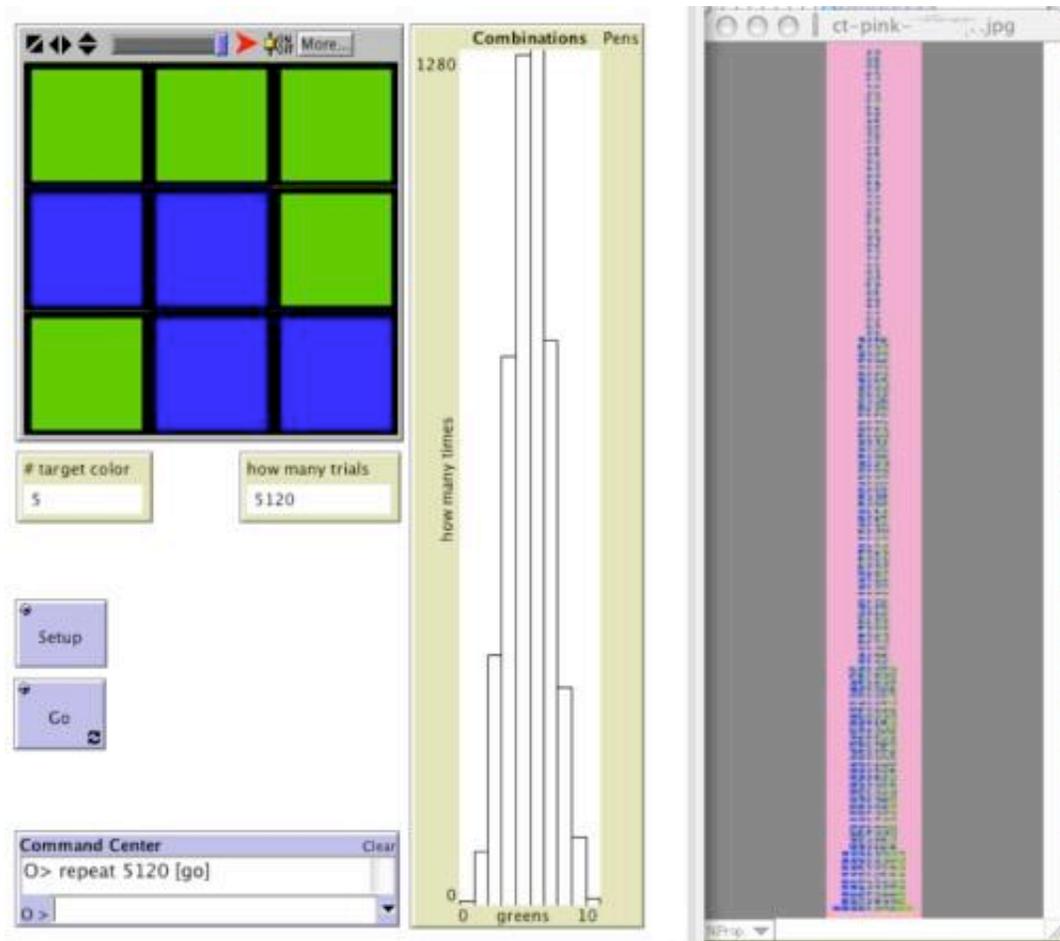
Figure 3. A combinations tower in the NetLogo probability experiment (the complete tower is on the left, and an enlarged fragment is on the right). The tower is the exhaustive combinatorial sample space of all 512 3-by-3 arrays in which each of the nine squares can be either green or blue.

1.4 Empirical Probability With Computer-Based Simulated Experiments

To produce the pictures in Figure 3 for this article we ran one of several interactive computer models for probability and statistics that we created in the NetLogo modeling environment ([Wilensky 1999](#)). These ProbLab models were used in the classroom to run empirical-probability experiments (see [Figure 4](#)). The experiments produced histograms that had the same overall shape as the combinations tower students had built and glued to the classroom wall. Note that whereas the combinations tower contains a single specimen of each of the 512 combinations, the empirical-probability experiments generate repetitions. For instance, the all-green 9-block could occur more than just once, just as ‘heads’ occurs many times when you flip a coin repeatedly. However, because each of the 9-blocks are equally likely to occur, the columns with repetitions roughly maintain their relative sizes as in the combinations tower. For instance, the two-green column grows to approximately four times the height of the one-green column that is also growing (see [Figure 4](#)). The empirically-generated columns may be taller because they reflect the results of many trials, but the more trials we run, the more these empirical histograms resemble in shape the sample space from which they are drawn. Thus, with respect to the empirical distribution, the combinations tower is the shape of things to come. That is, for a sufficiently large number of runs, the theoretical and empirical towers are proportionate to each other—they are the same shape only different in size. This stabilizing of the histogram on a “limit” illustrates the Law of Large Numbers. (All the simulations mentioned in this paper are available for free download as part of the NetLogo package at ccl.northwestern.edu/netlogo/. To run these experiments online without downloading NetLogo, go to ccl.northwester.edu/curriculum/ProbLab/)



After 25 trials After 100 trials After 200 trials After 512 trials After 5120 trials After 51200 trials
[Figure 4a](#) [Figure 4b](#) [Figure 4c](#) [Figure 4d](#) [Figure 4e](#) [Figure 4f](#)



[Figure 4g](#)

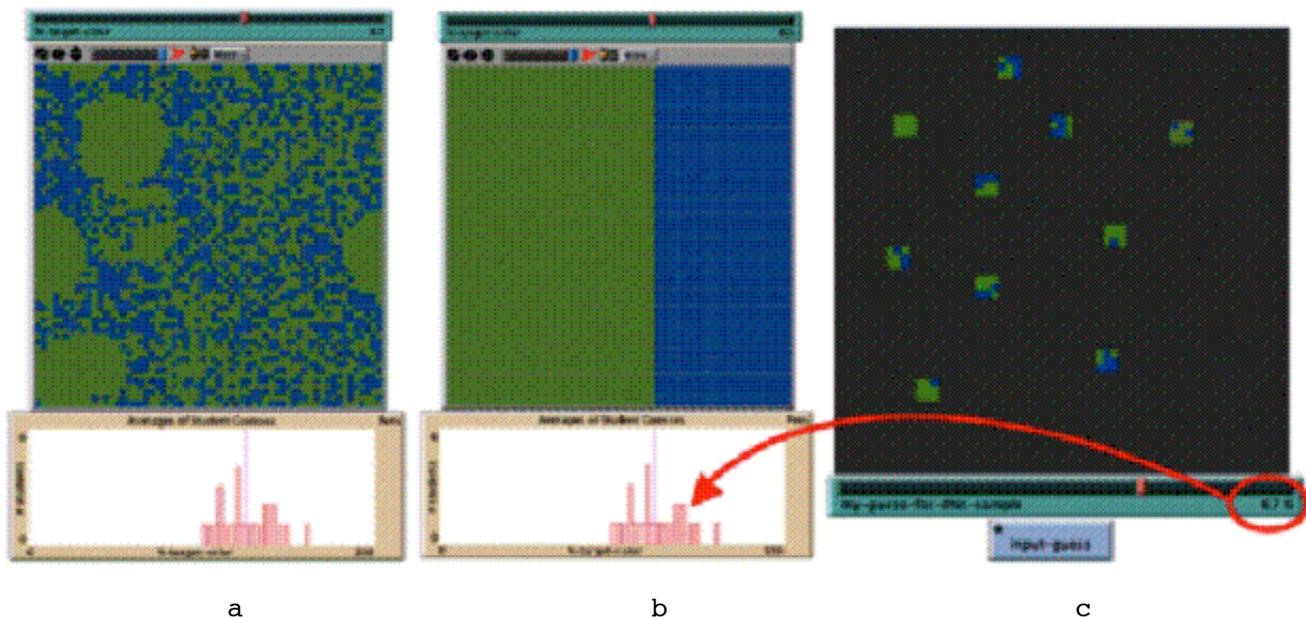
Figure 4. On top: In simulating an experiment in probability, a NetLogo interactive computer model generated 9-blocks randomly and plotted their cumulative distribution according to the number of green squares in each. As the simulation runs, the distribution tends in shape towards the sample space from which random samples are chosen, that is, the combinations tower. On the bottom: This is a screenshot from the teacher's computer interface that is projected onto the classroom screen. On the left is the NetLogo model that produces an occurrences distribution as it runs (empirical probability), and on the right is a picture of the combinations tower produced by another model and resembling the classroom combinations tower that students built (see also [Abrahamson \(in press\)](#)).

1.5 Learning Basic Statistics With S.A.M.P.L.E.R.

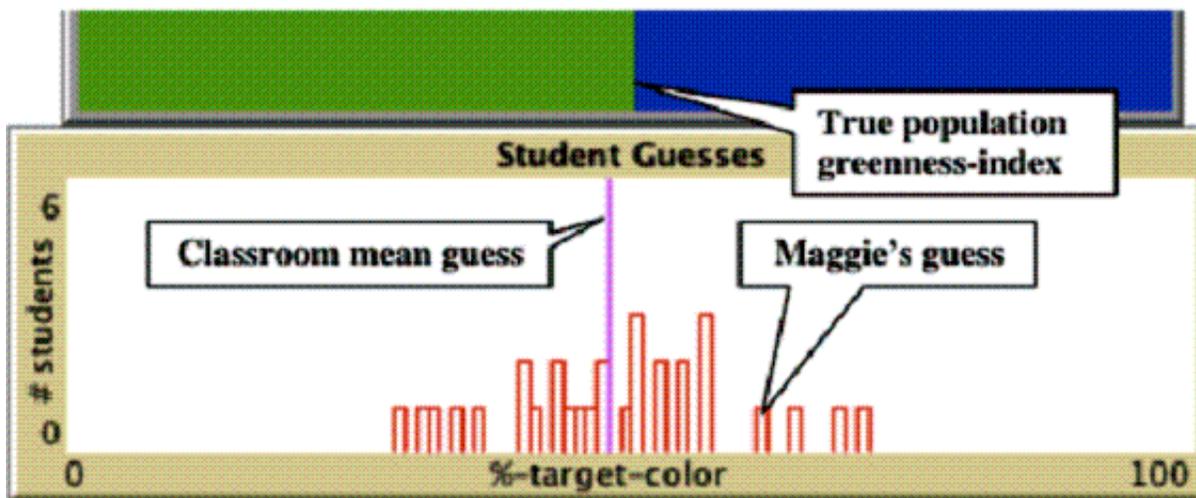
In this section we will discuss the S.A.M.P.L.E.R. design and describe a typical use of this design in a middle-school classroom.

1.5.1 S.A.M.P.L.E.R.: An Introduction

S.A.M.P.L.E.R., Statistics As Multi-Participant Learning-Environment Resource, is a participatory simulation activity ([Wilensky and Stroup 1999b](#)) implemented in the HubNet ([Wilensky and Stroup 1999a](#)) technological infrastructure, which extends NetLogo so as to enable facilitation of collaborative inquiry in a networked classroom. In S.A.M.P.L.E.R. (see [Figure 5](#)), students take individual samples from a population so as to determine a target property of this population. The “population” is a matrix of thousands of green or blue squares ([Figure 5a](#)) and the target property being measured is the population’s greenness, i.e., the proportion of green in the population. A feature of the activity is that population squares can be “organized”—all green to the left, all blue to the right ([Figure 5b](#)). This “organizing” indexes the proportion of greenness as a part-to-whole linear extension that maps onto scales that are both in a slider (above the population) and in a histogram of students’ collective guesses (below the population). The population can be set to bear an unknown random percentage of green squares, and the population can then be hidden (masked) so that information about the green/blue properties of the squares can be gleaned only through sampling. Students participate through clients (in the current version of S.A.M.P.L.E.R., these clients run on students’ personal computers). These clients are hooked up to the facilitator’s server. Students take individual samples from the population ([Figure 5c](#)), and analyze these samples so as to establish their best guess for the population’s target property, its greenness. Note that whereas all students sample from the same population, by default each student only sees their own samples, unless these are “pooled” on the server. Students input their individual guesses and these guesses are processed through the central server and displayed as a histogram on the server’s interface that is projected upon the classroom overhead screen (see fragment of this projection in [Figure 5d](#)).



[Figure 5abc](#)



d

[Figure 5d](#)

Figure 5. Selected features of the S.A.M.P.L.E.R. computer-based learning environment.

The projected histogram shows all student guesses and the classroom mean guess, and this histogram interfaces with the self-indexing green–blue population. Note the small gap ([Figure 5d](#), middle) between the classroom mean guess and the true population index. Because a classroom-full of students takes different samples from the same population, the histogram of collective student input typically approximates a normal distribution and the mean approximates the true value of the target property being measured. The students themselves constitute data points on the plot (“I am the 37” ... “So am I!” ... “Oh no ... who is the 81?!”). So students can reflect both on their individual guesses as compared to their classmates’ guesses and on the classroom guess as compared to the population’s true value of greenness. Such reflection and the discussion it stimulates are designed to foster opportunities for discussing and understanding typical distributions of sample means.

S.A.M.P.L.E.R. can constitute a standalone set of activities, yet the general framework of ProbLab is for students to participate in activities that interleave and juxtapose the statistics component, the theoretical-probability component, and the empirical-probability component. Activities are designed for the 9-block to play a pivotal role in students’ bridging between S.A.M.P.L.E.R. and the other pillars of ProbLab. The 9-block features in S.A.M.P.L.E.R. as samples of size 3 by 3. Students taking 3-by-3 samples from the S.A.M.P.L.E.R. population may construe the greenness of the population in terms of 9-blocks, and this interpretation may help students bridge from statistics to both theoretical and empirical probability, as follows.

Students may bridge between statistics and theoretical-probability by comparing between the S.A.M.P.L.E.R. population and the combinations tower. Specifically, students may construe a S.A.M.P.L.E.R. population as a collage from “the right side” (more green than blue) or “the left side” (more blue than green) of the combinations tower.

Students may bridge between statistics and empirical-probability using 9-block distributions: the act of

sampling a 9-block from the S.A.M.P.L.E.R. population is meaningfully related to generating a random 9-block, e.g., as in the 9-Blocks interactive computer model. Both in S.A.M.P.L.E.R. and in the 9-Block model, the user expects to receive a 9-block but does not know which 9-block will appear on the interface. So students may think of the S.A.M.P.L.E.R. population as a collection of many random 9-blocks. This may support students in developing and using sophisticated techniques for evaluating the greenness of the S.A.M.P.L.E.R. population. Specifically, students may attend to each 9-block sample individually, and learn to use histograms so as to record sample values as distributions. Otherwise, students often count up all the green little squares they have exposed and then divide this total by the total number of exposed squares, in order to determine the greenness of the population. Such a strategy, albeit effective for achieving the goal of evaluating the greenness of the population, misses out on a learning opportunity, because it does not make for mathematizing the variety of samples as a distribution—it “collapses” the variation, resulting in an impoverished notion of distribution. Therefore, bridging between probability-and-statistics activities is helpful not only for building a cohesive understanding of the entire unit but also for understanding each of the conceptual pillars of this unit (see also [Abrahamson \(2006\)](#)).

1.5.2 Learning with S.A.M.P.L.E.R.

Whereas participatory simulation activities (PSA ([Wilensky and Stroup 1999b](#))) may take many trajectories, depending on facilitators’ goals and learners’ age and interest, we have found it useful to describe “typical” implementations of our PSAs when we have a clear idea both of the age group and the broader curricular context (see HubNet participatory-simulation guides at ccl.northwestern.edu/netlogo/hubnet.html). Such descriptions have helped teachers, and in particular teachers who are new to networked classrooms, prepare for facilitating the PSA in their own classrooms. The following description is based on pilot studies with focus-groups and classrooms ([Abrahamson and Wilensky 2004b](#)).

The implementation of S.A.M.P.L.E.R. follows three stages: introduction (server only), student-led sampling and analysis (server only); and collaborative simulation (clients and server). Typically, the first two stages take between half an hour and an hour, depending on student age group. The third stage may take between one and three periods, depending on student engagement and the teacher’s flexibility in “weaving into” the PSA other ProbLab activities, such as NetLogo models, that may challenge students to reason carefully and thus deepen and enrich the discussion.

Introduction. The activity begins with the facilitator showing students a population of green and blue squares (the population is entirely exposed). Students offer their interpretations of what they are seeing. The teacher then asks students how green the population is, and students discuss the meaning of the question, offer intuitive responses, reflect on the diversity of responses in their classroom, articulate personal strategies, and develop more rigorous strategies and suggest how they could be implemented in the computer environment. The teacher facilitates the discussion by reminding students of mathematical content they had studied in the past that appears relevant to students’ intuitive strategies. In doing so, the teacher introduces mathematical vocabulary that will help students communicate during the activity. For instance, a student might say, “It’s too much to count all of the little squares—if only we could just look at one little place and decide with that,” to which the teacher may respond, “So you want to focus on just a sample of this entire population of squares—how should we decide what a good sample is that will allow us to make a calculated guess or predict the greenness of the entire population?”

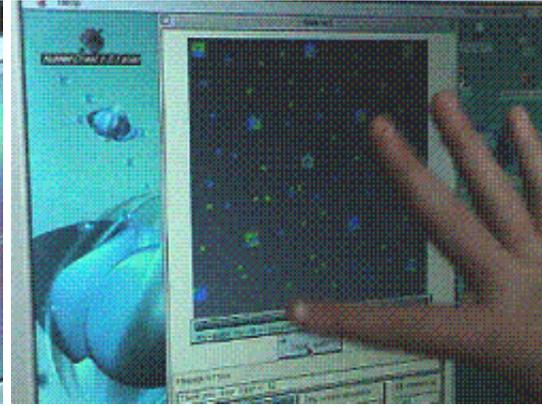
Student-led sampling and analysis. The teacher creates a new population that is not exposed. A student uses the teacher’s computer, which is functioning as the “server” of the activity, to take a single sample

from the population. To determine the size of this sample and its location in the population grid, the student–leader takes suggestions from classmates, asking individuals to warrant their suggestions. Once the sample is taken, by clicking with the mouse on a selected point in the population, students discuss the meaning of this sample in terms of the goal of determining the population’s greenness. For example, if a 5-by-5 sample has 4 green squares and 21 blue squares, students may want first to describe it mathematically, e.g., “The ratio is 4 to 25” (correct), and then draw conclusions from this sample, e.g., “There are 16 green squares on the whole screen, because $4/25$ is like $16/100$ ” (partially correct). Students then debate over the location and size of another sample, further discussion ensues based on this new sample, and then more samples are taken. The teacher encourages students to keep a record of the data and to draw conclusions from the accumulated data. For instance, let us assume that students have taken ten samples each of 25 squares and have received the following data, couched in terms of the number of green squares in each sample: 8, 4, 4, 9, 21, 6, 4, 8, 9, 7. What are we to do with these data? Sum them all up? Decide that the answer is “4,” because “4” occurred more than any other number? Ignore the “21,” because it does not fit with the others? Calculate the average - 8 - and state that 8% of the population is green? Perhaps we should conclude that, seeing as the samples are inconsistent, these data are useless? The teacher guides students towards effective procedures by recording all the ideas and then exposing the population and discussing with students which procedure appears to be yielding the best results over repeated trials.

Collaborative simulation. The teacher creates a new unexposed population and, through the server’s interface, enables students’ sampling functionalities. Students each take samples. The total number of little squares students may expose is limited by a “sampling allowance,” for instance a total of 125 squares, that the facilitator sets from the server. This allowance is “replenished” between rounds. To optimize the gain from their limited sampling allowance, students each strategize the size and number of their individual samples as well as the location of these samples on the population grid. [Figure 6](#) illustrates two different strategies students often use. One student (see [Figure 6a](#)) worked in the “few–big” strategy, spending the allowance mostly on a single location where the student took an 11-by-11 sample (a 121-block). Students who operate thus often say they are trying to create a reduced picture of the entire population. Some of these students choose to take the large sample from the center of the population (and not from a corner as in [Figure 6a](#)) and say that the center is the most representative location for the whole population. They also suggest that they can more readily calculate the proportion of green in their samples if they take just one sample and not many. Another student (see [Figure 6b](#)) worked in the “many–small” strategy, spending the sampling allowance by scattering samples of size 3-by-3 (9-blocks) and 1-by-1 (1-blocks) in a more-or-less uniform pattern across the population. Students operating thus often say they are trying to cover as much ground as possible, in case there is variance in the population that could not be found through a single large sample. Also, the “many–small” students are more likely than the “few–big” students to use averaging methods in analyzing their sampling data. Classroom discussions address individual techniques for maximizing the utility of the limited sampling resources and for making sense of the data.



[Figure 6a](#)



[Figure 6b](#)

Figure 6. Examples of student sampling strategies: “few-big” and “many-small.”

At the end of each round, students use a slider to indicate their guess for the population’s greenness, e.g., 83%, and press a button to input this guess to the server. A histogram that shows all students’ guesses is thus projected on the overhead screen. Often, this histogram approximates a bell shape. The teacher exposes the population and then “organizes” it so that the population’s true value of greenness is evident. Whereas individual students may be up to 20 or more percentile points off mark of the true value, the mean of the histogram—the “class guess”—is often less than 5% away. Moreover, often no student has input the value of the classroom mean guess—this mean is indeed only the guess of the classroom as a whole.

An optional feature of S.A.M.P.L.E.R. is that students begin each round with 100 personal “points.” When students input their guess, they also commit either to their personal guess or to the classroom mean guess. Once students have input their guesses, each student has some points deducted according to the error of the guess they had committed to. For instance, based on her samples, Maggie input “70%” and committed to her personal guess. Assuming the true value of greenness turns out to be 50%, Maggie will lose 20 points. But, assuming that the class’s mean guess is 55%, had Maggie committed to the class guess, she’d have lost only 5 points. The juxtaposition of personal and pooled accuracy often engenders a pivotal moment in the activity: as individuals, students each can view themselves as a single data point on the histogram, but as an aggregate, the classroom embodies a distribution. This identity tug-of-war, “me vs. classroom,” that is stoked by personal stakes in the guessing game and by social dynamics around this game, is designed to provide opportunities for students to ground the ideas of distribution and mean.

Once the classroom guesses have been plotted as a histogram and the true value of greenness has been exposed, volunteer students go up to the front of the classroom, explain the histogram, analyze the accuracy of the classroom guess, and respond to their classmates’ questions. In particular, students share their personal sampling- and data-analysis strategies in a collaborative attempt to improve on the accuracy of the classroom mean guess on a subsequent round.

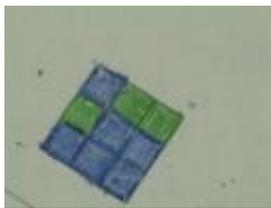
Following several practice rounds, the facilitator may challenge students by decreasing the sampling allowance so that students each have limited personal information about the population. Some local as well as classroom-level spontaneous conversation may emerge, through which students coordinate their sampling so as to maximize the total exposed area in the population (because it would be redundant to take multiple samples from the same location). If students conclude that it is better, individually, to “go with the group guess,” should the group somehow collaborate to ensure higher accuracy? Some students believe that, once

a new population is created and students have taken samples, it is better first to discuss their estimations and then input guesses rather than first to input their guesses and then discuss the distribution. These students argue that by first discussing, the group can decide on a single guess and, thus, ostensibly, achieve higher accuracy. This strategy, which obviates the range and variance of the distribution, affords an opportunity to discuss properties of the distribution and recontextualize the rote procedures of calculating a mean.

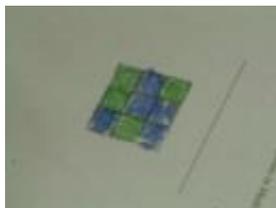
With the description of S.A.M.P.L.E.R., we have concluded the introduction of the design of ProbLab. We will now turn to examine data from an implementation of ProbLab in the second author's two middle-school classrooms.

2. Classroom Data

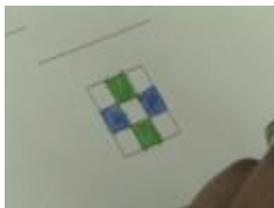
The combinations tower lessons were co-taught by the first two authors in two 6-grade classrooms (in a 6–8 middle school with 52% minority students and 36% students eligible for free/reduced lunch). Initially, before our 6-grade students had been told they would be building the combinations tower, we gave them worksheets with blank 9-blocks and asked the students to fill in the squares with either green or blue crayons (see [Figure 7](#)). We then asked students to compare their patterns to those created by their classmates. Next, we asked how many different green–blue combinations there can be at all. We wrote students' guesses on the blackboard and asked each student to explain their thinking.



[Figure 7a](#)



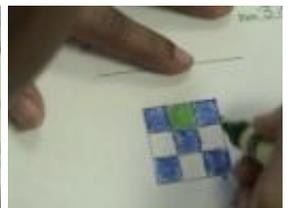
[Figure 7b](#)



[Figure 7c](#)



[Figure 7d](#)



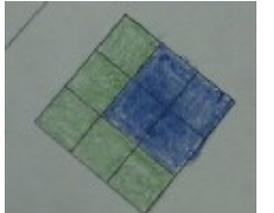
[Figure 7e](#)



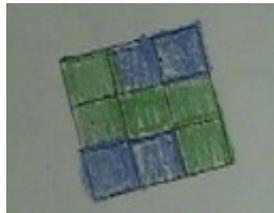
[Figure 7f](#)



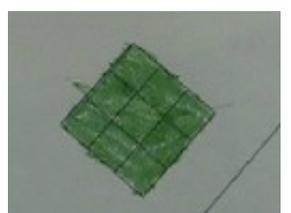
[Figure 7g](#)



[Figure 7h](#)



[Figure 7i](#)



[Figure 7j](#)

Figure 7. Variation in individual student work on the first task of creating a green-blue pattern.

Next, we gave students worksheets with thirty-two blank 9-blocks, and asked the students to fill in as many different combinations as they are able. After some individual work (see [Figure 8](#)), students came to realize that there are many more possible combinations than they had initially estimated.

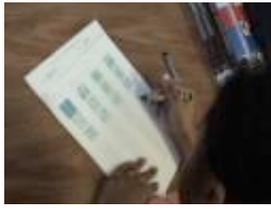
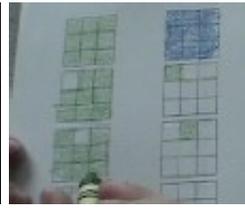
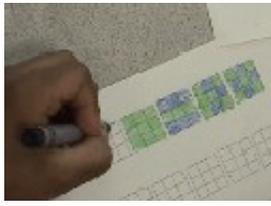
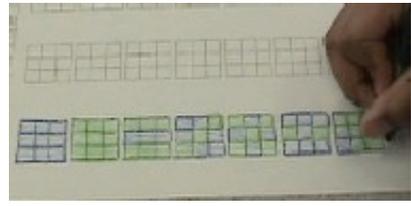
[Figure 8a](#)[Figure 8b](#)[Figure 8c](#)[Figure 8d](#)[Figure 8e](#)[Figure 8f](#)[Figure 8g](#)[Figure 8h](#)

Figure 8. Students work on creating different combinations of the green-blue 9-block. They use personal methods that range and develop from explorative to rigorous.

We showed students a NetLogo model, “9-Blocks,” model that randomly created blue–green 9-blocks in succession (see [Figure 9](#), for a set of screenshots that show a fragment of the interface with the virtual 9-block). Students commented that neither is the computer working according to any method nor is it keeping track of whether or not it is repeating guesses. These observations stimulated students to discuss potential methods for rigorously determining all the different combinations. This discussion also engaged students who had not been methodical in their individual work yet came to appreciate the need for a logical–mathematical procedure (algorithm).

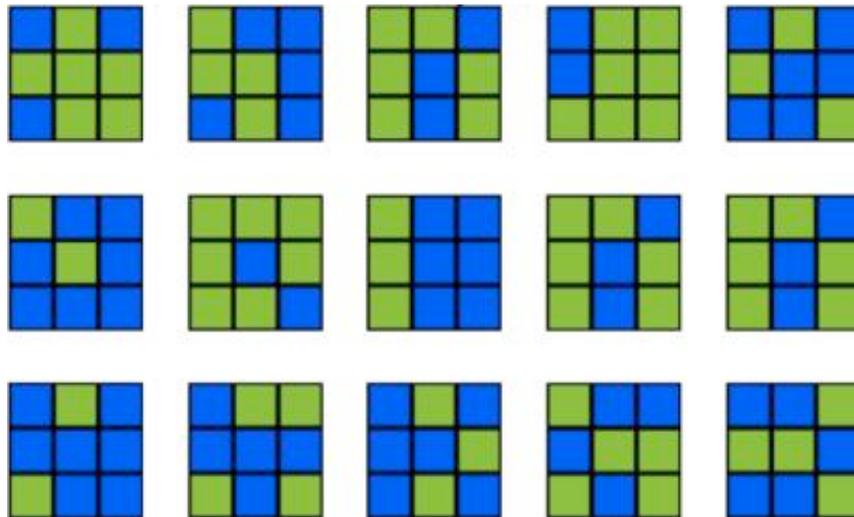
[Figure 9](#)

Figure 9. These are fifteen separate screenshots from the NetLogo model Stochastic Patchwork. The model generates such random combinations successively. The user can control the number of squares in the sample as well as the speed of the experiment and other parameters. In this particular run there happened to be a

single repetition (the 3rd, 9th, and the 10th samples).

2.1 Students' Initial Discoveries

Many of the students used the table format of our worksheet to organize their mathematical inquiry. The typically higher-achieving students discovered methods for determining all the combinations with two green squares. They developed vocabulary for communicating and debating their methods. For instance, they named one square the “anchor” and the other the “mover.” The anchor and mover are temporary names for squares in the 9-block that help keep track of the search for all combinations. For instance, the anchor is initially the top-left square, and the mover visits in each of the remaining eight squares. Following, the anchor moves one square, and the mover visits the remaining seven squares. Thus, the total number of combinations with two green squares is $8 + 7 + 6 + 5 + 4 + 3 + 2 + 1 = 36$. A couple of students explained that they had initially thought that when the anchor was in the second square, the mover could be in the first, but then they found that such a strategy would include repetitions. We encouraged student to notice that if we were to keep all repetitions, we would have $9 * 8$ combinations, and then we could divide by 2 to receive 36. Thus, we inadvertently discussed an example of the formula $n(n - 1) / 2$ that gives the sum of the series $1 + 2 + 3 + \dots + n$. Note that this is a case of the formula for finding the binomial coefficients, in which $n = 2$, $N = 9$, and p and q are 0.5.

Many students noticed how the total number of green and blue squares in the 9-block is complementary—a 7-green block is the same as a 2-blue block—and so one need not create both of these but only one of them, because they are essentially the same class of 9-blocks. One strategy students used was to create a combination and then reverse the green and blue colors. This work helped students realize that the number of different combinations with, say, two green squares is the same as the number of combinations with two blue squares. This means that the distribution of combinations itself is symmetrical (as in the combinations tower). Students who had made these discoveries became leaders who presented and explained their ideas to the whole classroom. [Figure 10](#) summarizes some of the students' insights. The figure shows how green and blue are complementary. The bottom row, “combinations,” represents the number of different 9-blocks in each column. At this point, at the end of the first double period, only three values have been discovered (reading from the bottom-right corner, and moving to the left): 1 combination with 9 green squares (or 0 blue); 9 different combinations each with 8 green squares (or 1 blue), and 36 different combinations each with 7 green squares (or 2 blue). The ‘AM’ and ‘PM’ captions designate which part of the construction work will be done by each of Ms. Janusz's two classrooms, the morning class and afternoon class. It was our idea that the two classrooms collaborate, and student leaders were glad to organize this collaboration because several of them were personal friends with students in the other classroom.

	AM					PM				
# Green	0	1	2	3	4	5	6	7	8	9
# Blue	9	8	7	6	5	4	3	2	1	0
Combinations										
	3691									

[Figure 10](#)

Figure 10. A table of combinatorial analysis that summarizes students' discoveries on the first day at the end of the double period. Students found that there are 36 different combinations with seven green squares (or two blue squares; see bottom right corner).

Note, in Figure 10, that the PM classroom has found three values for the table – 36, 9, and 1 – whereas the AM classroom has not filled in any. This apparent advantage of the PM group is by and large an artifact of the experimental design. The first two authors were implementing this experimental design for the first time, so the afternoon classroom periodically enjoyed improved facilitation, including better organization in using available learning tools to support classroom collaboration and discussion. As it turned out, this advantage was helpful, because in the AM classroom there were several more advanced students than in the PM classroom.

2.2 Engineering and Building the 9-Blocks

Students realized that there are many combinations and that in order to create all these combinations, they must address the time limitations. So it became apparent to the students that they would need to pool their resources so as to produce the entire sample space within the allotted time. Students said they would need to share their methods for determining all the different combinations as well as the labor of creating these paper-and-crayon combinations. Also, students understood that such engineering requires careful group management to increase the efficiency of the group. We guided the students to arrive at the particular management strategy that would be conducive to organizing the production of combinations according to the number of green squares in each 9-block. Initially, students suggested to divide the class into groups that would each be responsible for creating a different type of 9-block: one group was to create all 9-blocks with no green squares, another group was to create all 9-blocks with one green square, yet another group was to create all 9-blocks with two green squares, etc. Based on their earlier work, the groups quickly realized that

the task was unfairly distributed among groups. It seemed as though the more green squares in a 9-block, the more students are required to create that set of 9 blocks. So the number of students on task for each category of 9-blocks and the total student time spent on each category was roughly proportionate to the number of 9-blocks in each category. The classroom was thus distributed into task-teams of relative sizes that corresponded to the sample space they were building!

Several students voluntarily led the classroom work. These were students who typically were more enthusiastic about learning mathematics. After the lesson was over, these students spontaneously approached us to discuss the activities. In the subsequent lesson, these ‘student leaders’ took on the following responsibilities:

- Design the collaborative engineering task, including the distribution of students between construction groups, with a sensitivity to their classmates’ mathematical prowess
- Assign tasks to students and explain these tasks to them
- Circle in the classroom to supervise the production of the combinations
- Manage the construction of the combinations tower

We introduced the combinations tower by challenging students to engineer a display that would help us to compare easily between the subgroups of 9-blocks that they had created (with zero-green, one-green, two-green, etc.). Also, we advised students that their display should communicate to non-participants their discovery of the symmetry of the distribution.

2.3 The Emergent Social–Mathematical Space

On the second day of the implementation, students continued working in groups on producing all the different 9-block combinations. Over the course of the work, students self organized into different roles, in and between groups, according to the following types of emergent expertise: (1) coordinators who managed the overall classroom collaboration, networked, and circled between groups to teach the combinatorial-analysis method and support its implementation; (2) “number crunchers” who relied primarily on theoretical analysis of the problems to find the number of different combinations (they were less enthusiastic actually to build these combinations); (3) designers (engineers) who devised visual methods (algorithms) for finding all the combinations without repetitions; (4) implementers who carried out the designed methods by indicating with a pencil on the 9-block squares and sometimes acted as local group leaders; (5) producers who filled in with crayons according to the implementers’ pencil indications; (6) checkers (quality-assurance experts) who searched for repeated combinations and crossed them out and made sure that the 9-blocks were glued in the correct orientation into the combinations tower; and (7) constructors who cut the prepared and approved 9-blocks out of the worksheets and glued them onto the larger paper canvas into the appropriate columns. That students each engaged in a role that suited their mathematical capability is an example of tradeoffs inherent in collaborative projects (for a discussion of this *stratified learning zone*, see [Abrahamson and Wilensky \(2005\)](#)).

2.4 Student Reasoning

Students explored and connected between various methods for determining and creating all of the different 9-block combinations. We supported students in solidifying their understandings by asking them to communicate their ideas clearly to their group mates. Students responded by being more articulate. Following are some of the methods students developed and presented, according to the type of reasoning

involved:

2.4.1 Induction

A single square (a 1-block) has 2 possible combinations (green or blue); a block of two squares has 4 combinations (green–green, green–blue, blue–green, and blue–blue), and a block of three squares has 8 combinations. So for every square we add, the number of combinations doubles. Thus, for nine squares we get 512 combinations (see also [Piaget and Inhelder \(1969\)](#)), on children’s development from exploration to computation of combinations). We showed students the exponential function relating squares and combinations, that is $2^1 = 2$; $2^2 = 4$; $2^3 = 8$; $2^4 = 16$, etc.

2.4.2 Deduction

Students had found that there is 1 combination with no green squares (only blue squares), 1 combination with no blue squares (only green squares), 9 different combinations with exactly one green square (and eight blue squares), 9 different combinations with exactly one blue square (and eight green squares), 36 combinations with exactly two green squares (and seven blue squares), and 36 combinations with exactly two blue squares (and seven green squares; see [Figure 11](#)). So students knew that the distribution is symmetrical, and they also knew of a total 92 different combinations. Classroom discussion then proceeded as follows. If there is a total of 512 different combinations, then the number of different combinations that we have not accounted for is $512 - 92 = 420$. We can divide that by 2 to get a total of 210 unaccounted squares in the three-green and four-green groups. If we find out how many different combinations there are in the three-green group, we will know the number of combinations in the four-green group. We will also know the number of combinations in the three-blue group and the four-blue group.

0	1	2	3	4	5	6	7	8	9
9	8	7	6	5	4	3	2	1	0
1	9	36	84	126	126	84	36	9	1

[Figure 11a](#)



[Figure 11b](#)

Figure 11. Student achievement on the second day is seen in this figure: on the left is the distribution table with all the values filled in correctly and on the right is the combination tower with three complete columns on each side and four central columns yet to be assembled and built.

2.4.3 Recursion (spatial algorithm)

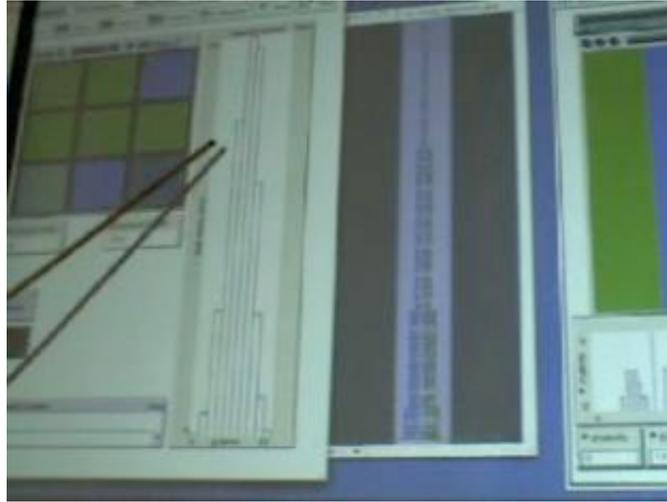
Students had found that the anchor–mover method works for the two-green group. For the three-green group, they added a “super anchor” and applied to it the anchor-mover method. For the four-green group, they then added yet a “super-duper anchor.”

2.4.4 Visual configurations

As an alternative to the anchor–mover method, a team of “designer” students tried to find all the combinations in the four-blue group by: (a) determining possible configurations of three blue squares, such as an “L,” a “diagonal,” a column or a row, etc.; (b) counting up the number of possible locations of these configuration in the 9-block; and (c) multiplying the total number of locations by 6, because the fourth blue square could be anywhere within the remaining six squares. In doing so, the students drew on the blackboard icons representing the spatial primitives they had agreed upon and named each primitive, while their teammates searched for these elements in the array of 9-blocks in the worksheets. (This method proved ultimately problematic due to duplications, e.g. a row with a fourth square anywhere on a row below it could also be construed as an “L” with a fourth square.) Another heuristic students explored was to compose a 9-block and then create its symmetrical, rotated, and symmetrically rotated 9-blocks (so there were different classes of 9-blocks in terms of the number of possibilities they could generate, with some original 9-blocks resulting in a total of eight different combinations).

The vocabulary that had been developed on the previous day, for instance “anchor” and “mover,” which students had developed for finding combinations with two green squares, was adopted by many students in both classrooms. Also, students elaborated on this vocabulary so as to accommodate the more complex problems they were addressing. Thus, words served as tools for students to communicate in engineering and constructing the combinations tower.

On their third day of working on the combinations tower, Ms. Janusz’s students from the morning and afternoon classes completed the tower. In the last 10 minutes in both of the morning and afternoon classes, we used this tower to help students relate theoretical and empirical probability (see [Figure 12](#)). One student, Emma, who had been quite reticent during the unit made the following observation in comparing the combinations tower (theoretical probability) and a NetLogo model that produced random 9-blocks and plotted their cumulative distribution (empirical probability; see also [Figure 4](#)).



[Figure 12](#)

Figure 12. A student explaining why a probability experiment (on the left) produces a histogram that resembles the representation produced through combinatorial analysis (in the center).

We had asked the classroom why it is that the histogram “grows” to resemble the tower. Specifically, dovetailing a student comment, we asked why it is that the 4-green column in the probability experiment was taller than other columns. Emma said:

“Maybe because there’s more of that kind of combination. Just basically, because if there’s 512 different combinations, and we know that there’s more [possible combinations] in the middle columns, [then] even though there’re duplicates, there’s still going to be more combinations in the middle columns. [The student is now using a pointer to explain what the class is watching on the screen] Even though these patterns [in the empirical live run, on left] may have duplicates in this [in the combinations tower, center] it’s still counting all the patterns, so it’s going to have the same shape.... It’s going to be the same shape, because it’s basically the same thing. Because in the world there are more patterns of these than there are of the other ones.”

Several of Emma’s classmates expressed similar budding understandings.

2.4.5 Sampling strategies and initial understandings

By way of introducing the S.A.M.P.L.E.R. activity, the teacher showed a revealed population and asked students for their interpretations. Students’ interpretations ranged from fanciful to “mathematical.” Fanciful interpretations included “Superman,” “a mushroom house,” “a man with a shield,” “a messed-up face,” a “teddy bear with eyes,” “a mountain,” and “an elephant with a nose.” Mathematical interpretations used the 9-block as a reference point. For instance, students said that the population could be thought of as a collection of many 9-blocks. Students wondered how many 9-blocks might fit into the population. This stimulated a discussion of methods for determining the number of squares in the population. Thinking in a different direction, one student suggested that we look at the entire population as a single “1000-block,” and

that what we were looking at was just one of many different combinations of this 1000-block. This interpretation appears to apply ideas that arose from the combinatorial-analysis and empirical-probability activities to the entire population of squares.

Next, to introduce the sampling activity feature, the facilitator used a population that was entirely revealed but for which the greenness value was not disclosed. The facilitator asked the students how one could determine the greenness of the population. Students said that, in principle, one could count up all the green squares in the population and divide this number by the total number of squares in the population. However, students said, there are too many little squares to count, making this strategy unfeasible. Other students suggested that it might be useful to focus on a single area of the population and count up the green squares in it. The facilitator reiterated this idea, calling that area a “sample.” The question on the table then became, “If we could only take a single sample, where should we take it from?” The following transcription illustrates classroom discussion about sampling.

Student 1: It would be better if there were a way to get a random spot. [for the sample]

Researcher: A random spot?

St. 1: Yeah, because if you chose somewhere, you might think, “Mmm, this one has a lot of green, let’s do it there.”

Res: But what if *randomly* the computer gives me a place with a lot of green or a lot of blue?

St. 1: Well, then that’s what you’ve got to guess on.

St. 2: [You should put the sample] in the middle, a little higher ... it seems a little sort of balanced.

St. 1: But that’s just what I’m saying. If you try to find something balanced, it’s going to be around 50% no matter what.

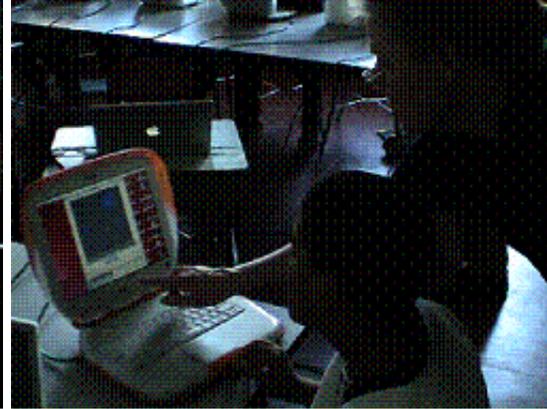
These students’ exchange reflects a pivotal quandary of statistics—is the sample sufficiently representative of the population, and what measures can we take to ensure that it is? A feature of the design that supported this conversation was that the facilitator could toggle between a view of the whole population and a view of different samples. Thus, students could gauge whether various suggested sample sets were sufficiently representative of the population. Most students did not use proportion-based mathematical vocabulary, possibly because they were not fluent in its application to novel situations. Yet, the visualization features of the learning environment enabled these students to communicate about proportionality qualitatively.

The lesson continued with students working on their individual computers. Students took samples from the population, inputted their guesses to the server, and examined results once the population and its true greenness value were revealed. The teacher worked with individual students as they participated in these activities. In [Figure 13a](#), the teacher is working with one of the students she had listed as high achieving in mathematics. They are comparing the student’s guess for the population’s greenness with the population’s true greenness value. In particular, the student is showing the teacher that she had guessed correctly—the green–blue partition in the population is precisely where the student had indicated it would be. The student

explains to the teacher her sampling strategy. In [Figure 13b](#), the teacher is working with one of the students she had listed as low achieving in mathematics. The student had taken samples from the population and had input a guess that did not seem to reflect all the samples he had taken. The teacher is discussing with the student whether it would help for him to consider all samples the in determining the greenness of the population. These classroom data demonstrate both that the S.A.M.P.L.E.R. activity enables immediate feedback to the teacher and helps the teacher elicit a wide range of student difficulties, which she can then address in classroom discussions. Also, these data demonstrate one way that PSA integrate group- and individual work: the framework of the activity is collaborative, but to participate successfully in this collaboration, students must each achieve an understanding of the activity.



[Figure 13a](#)



[Figure 13b](#)

Figure 13. During student work, the teacher has opportunities to work with each student.

2.5 Post-Test Results

The implementation in Ms. Janusz's classroom spanned over one week (Monday—Friday), with the first three days dedicated primarily to constructing the combinations tower and the remainder of the time spent on working with S.A.M.P.L.E.R. Following, we report on findings from an analysis of student response to several post-intervention questionnaire items. The AM classroom scored slightly higher than the PM class on these tests, yet that advantage may have resulted from the higher number of advanced students in the AM classroom.

2.5.1 Combinatorial analysis.

In total, students worked approximately two-and-a-half 80-min. periods on building the combinations tower and viewing computer models that generated random 9-blocks. Following this mini-unit, students completed a questionnaire that included feedback questions and some content questions. Over half of the students completed correctly the following item:

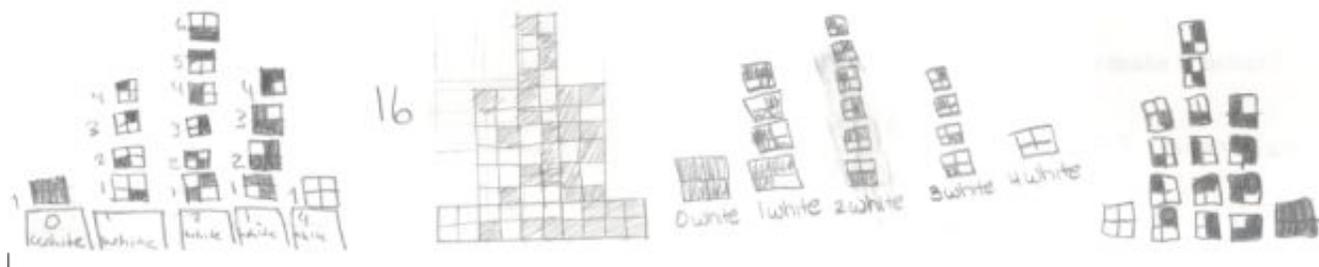
This is a 4-block. It is empty.





- Make a combinations tower of all the different black-and-white 4 blocks.
- If a computer were making 4-blocks randomly, what is the chance of getting a 4-block with exactly 3 black squares?

Of the students who did not respond correctly, about a half were not careful enough in constructing the 4-block combination tower, so either they left out or duplicated one or more blocks (see [Figure 14](#), for examples of constructions that led to correct responses). Bearing in mind the 12th grade students 3% success on a comparable item (NCES, 2004b), our 6th grade students' performance of over 50% correct indicates that this mini-unit may constitute a contribution to helping students build meaning for probability.



[Figure 14](#)

Figure 14. Four examples of student constructions of the 4-block combinatorial sample space. Over half of the students could build these combination towers and solve correctly a probability question concerning the chance of producing a 4-block with exactly 3 black squares.

2.5.2 Statistical analysis.

Students worked a total of about 2.5 double periods on S.A.M.P.L.E.R. Students' responses on the post-intervention questionnaire revealed a wide range in classroom experiences in the unit. Responding to an item requesting their favorite sampling strategy, many students said they enjoyed spreading their samples all over the screen and then counting up the total number of green squares, dividing this number by the total number of exposed squares, and calculating this quotient as a percentage. Of these students, some thought that it is better to take single-square samples so as to maximize the spread of squares. Other students said that distributing their samples "randomly" was a better strategy as compared to distributing them systematically. Many students thought it is highly efficient to divide the sampling task between many students by allocating specific sampling areas to specific students—this strategy, they wrote, maximizes the total exposed squares.

Another item asked students whether it is better to commit to one's own guess or to commit to the group guess. That is, which of these two strategies ensures better long-term results? Students' answers varied, and they depended on the students' mathematical ability. High-achieving students preferred going alone, unless they were very unsure of themselves, whereas lower-achieving students preferred to trust the group guess. So the lower-achieving students were those who believed that the compiled guess is a more accurate

measure of the statistical data as compared to an individual guess. This finding is somewhat counter-intuitive. One might expect that the higher-achieving students and not the lower-achieving students would be those who gain this mathematical insight. Possibly, the higher-achieving students are those who more often suffered from their classmates' "wayward guesses," i.e. off-mark input that resulted from incorrect analysis and not from "extreme" sample. So the accuracy of students' individual guesses resulted both from a random factor—the specific samples each student exposed—and from a skill factor, students' individual mathematical competency reflected in their ability to calculate a percentage.

In their written responses, all students referred in one way or another to the distribution and range of the guesses, couching these in terms of 'left,' 'right,' average, and balancing ("it evens out"). We interpret this finding as indicating that the S.A.M.P.L.E.R. activities created a shared classroom artifact that carried shared meanings, experiences, and vocabulary. Such shared mathematical images could serve as helpful anchors in future classroom discussions.

Yet another item asked students whether one should first input a guess and only then discuss the input or first discuss and then guess. Many students thought that discussing first might either confuse you or bias the group guess—that a wider distribution guaranteed more accuracy of the classroom group guess. We interpret this finding as indicating that students experienced how an aggregation of random outcomes can nevertheless effect higher accuracy than would a "centralized command" (see also [Wilensky \(1997, 2001\)](#); [Surowiecki \(2004\)](#)).

Finally, students varied in what they considered to be a "good guess." Some students were happy to be several percentage points off the true value, whereas other students were more critical of their guesses (for a more detailed report on students' spontaneous sampling strategies, see ([Abrahamson and Wilensky 2004b](#))).

3. Conclusion and Future Work

High-school, college, and even graduate-school students struggle with making sense of probability and statistics. If 6th-grade students participating in our design have opportunities to relate to deep mathematical ideas of this domain, even if qualitatively, the design rationale and materials may hold promise for fostering learning trajectories into high school. Furthermore, it may be that ample classroom time should be allocated to difficult construction tasks such as collaboratively building a 9-blocks combinations tower. Only through such immersion can students create a mathematical "world," if to use Emma's term, that gives sense to mathematical concepts (see ([Abrahamson and Wilensky 2004c](#)) for further classroom episodes of students "re-inventing" statistics).

Our design of curricular material for technology-assisted middle-school mathematics-and-science classrooms is an ongoing project that is constantly informed by implementations in classrooms. Students' engagement in our activities—their high levels of participation, excitement, and feedback comments—encourage us to improve these activities and research them in more classrooms and with more teachers. One direction that we find promising is using the combinations tower activity so as to help students build meaning for the statistical concept 'normal distribution' (bell-shaped curve). We will now explain the rationale of this future work.

3.1 Using the Combination Towers to Understand Science

Teachers both of science and mathematics often wonder how they should help their students see and use the relations between these disciplines. This is especially the case for such teachers, as the second author, who teach the same group of students both science and mathematics. Often, science teachers use mathematical tools to make sense of data, and mathematics teachers use scientific context to demonstrate the use of mathematical knowledge. Our unit in probability may help teachers foreground a deep connection between the disciplines of science and mathematics by unlocking a riddle that is sometimes left as a given fact and overlooked as a potentially powerful learning experience.

We are all familiar with the bell-shaped ('normal') curve that characterizes many phenomena in science, biology, social sciences, evidence-based medicine, etc. For instance, if we were to measure the heights of all 6th-grade female students in the U.S. and plot these as a histogram, this histogram would be bell shaped. But why do these phenomena fall into this distribution? How can we make sense of this? (see also [Wilensky \(1997\)](#), for an analysis and meta-design solutions for students' difficulty with this concept). The combinations tower may serve as a clue or 'model' for beginning to tackle this puzzle.

For the sake of clarity, let us assume a grossly simplified scientific model that may then serve as a conceptual basis for understanding real phenomena that are complex. That is, if we accept this model as basically sound, we can then examine each of its specified assumptions so as to evaluate whether, how, and why the model falls short from representing the more complex reality. We would then explore how we may adapt and enrich this model so as to make it more general without losing its basic coherence. Such adaptation will possibly touch upon profound scientific ideas, because the model would give students a handle for articulating their intuitions mathematically.

In this basic mathematical–scientific explanatory model of height distribution in a large population:

- Only nine variable factors, genetic and environmental, contribute to determining a person's height, such as 'parents' height,' 'nutrition,' 'climate,' etc. These nine variables map onto the nine squares in the 9-block.
- Each person is randomly assigned values for each of the nine variables.
- Instead of looking at a range of values per variable factor, we will assume that each variable can only be either a 'positive' (green) or a 'negative' (blue).
- Each variable value has an equal likelihood (0.5 probability) for positive or negative, e.g., 'nutrition' could equally be either 'good' or 'bad.'
- Each of the nine factors contributes equally towards determining a person's height, e.g., 'good nutrition' would be as powerful as 'tall parents.'
- The more positive values a person has, the taller this person is, along some interval scale.
- The variables do not interact. For example, the contribution of a positive value for 'climate' is not dependent or affected by the value of any of the other variable properties.
- The range of the distribution of heights is limited both on the minimum and maximum (there is a "cutoff" standard deviation from the mean).

Given the above assumptions, the combinations tower is the sample space of all combinations of the nine variable factors contributing to a person's height, and these combinations are organized by "height groups" from 'shortest students' (on left) to 'tallest students' (on right). The leftmost 'shortest students' column and the rightmost 'tallest students' column each hold a single combination, the no-positive and the all-positive, respectively. In between, there are more combinations that yield a count of 4 or 5 positives as compared to, say, 3 or 6 positives, and there are more combinations that yield a count of 3 or 6 positives, as compared to

2 or 7, etc. Yet note that each of the 512 combinations is equally likely to occur. Thus, the bell curve can be understood as a combinatorial sample space of a cluster of variables that has been detected as contributing to a property of an observable phenomenon. Each variable is independent of the others, but as a cluster of variables that inform a property of a phenomenon, these nine variables are co-dependent, and these co-dependencies create the bell-shaped combinatorial distribution. This is why instances of a phenomenon are often distributed such that there are more “average” incidents. For example, there are more people of average height than there are short people or tall people.

By way of demonstrating how this model may be complexified, consider the dichotomous variable of each square in the generic model. If we were to increase the space of possible outcomes to 3, the combinations tower would grow to encompass $3^9 = 19,683$ different possibilities. The shape of this tower would be closer to a bell shape. If we were to modify the relative likelihoods or weight of individual squares or if we were to introduce causal contingencies between squares within the 9-block, we might affect the shape of the distribution. Implementing this model as a computer-based interactive simulation would enable us to readily explore the parameter space, tinker with the procedures underlying the emergent distribution, and receive immediate feedback in the form of mathematical representations. By way of comparing these simulated experiments to information from scientific and statistics resources, we can evaluate the explanatory power of our model and iteratively modify the model toward a better fit with the data.

3.2 Future Work

The development of ProbLab continues with research of student learning as they interact with computer-based models designed as environments for exploring central ideas of the domain. To interact with some of these recent models, go to ccl.northwestern.edu/curriculum/ProbLab/ and/or download the free NetLogo package, which includes many of these models, at ccl.northwestern.edu/netlogo/. Specifically, we wish to achieve a better understanding of whether and how the bridging tools foster deep understandings of the material and whether there are limitations to a design that encourages students to constantly compare and contrast topics that have traditionally been taught in separate units. We are in particular interested to evaluate whether middle-school students' insights can serve them as a basis for sustaining a sense of understanding as they appropriate formal mathematical expressions, such as the binomial function, and as they generalize their knowledge as tools for modeling and solving further exemplars related to the study of probability and statistics.

Acknowledgements

Research on ProbLab is sponsored by the NSF ROLE grant 0126227. The authors wish to thank the JSE reviewers for their insightful comments that helped us hone an articulation of the design rationale.

References

Abrahamson, D. (2004), “Keeping Meaning in Proportion: The Multiplication Table as a Case of Pedagogical Bridging Tools,” Unpublished doctoral dissertation, Northwestern University, Evanston, IL.

Abrahamson, D. (in press), “The Shape of Things to Come: The Computational Pictograph as a Bridge from

Combinatorial Space to Outcome Distribution,” *International Journal of Computers for Mathematics Learning*.

Abrahamson, D. (2006), “Bottom-up Stats: Toward an Agent-Based “Unified” Probability and Statistics,” in *Small Steps for Agents ... Giant Steps for Students?*, D. Abrahamson (Organizer), W. Wilensky (Chair), and M. Eisenberg (Discussant). Symposium conducted at the annual meeting of the American Educational Research Association, San Francisco, CA.

Abrahamson, D., Blikstein, P., Lamberty, K. K., and Wilensky, U. (2005). “Mixed-media learning environments,” in *Proceedings of the Fourth International Conference for Interaction Design and Children (IDC 2005)*, eds. M. Eisenberg and A. Eisenberg, Boulder, Colorado: IDC.

Abrahamson, D., and Wilensky, U. (2002), “ProbLab,” The Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.
ccl.northwestern.edu/curriculum/ProbLab/

Abrahamson, D., and Wilensky, U. (2004a), “ProbLab: A Computer-Supported Unit in Probability and Statistics,” in *Proceedings of the 28th Annual Meeting of the International Group for the Psychology of Mathematics Education Volume 1*, eds. M. J. Hoines and A. B. Fuglestad, Bergen, Norway: Bergen University College, p. 369.

Abrahamson, D., and Wilensky, U. (2004b), “S.A.M.P.L.E.R.: Collaborative Interactive Computer-Based Statistics Learning Environment,” in *Proceedings of the 10th International Congress on Mathematical Education*, ed. M. Niss, Copenhagen, Denmark.

Abrahamson, D., and Wilensky, U. (2004c), “S.A.M.P.L.E.R.: Statistics as Multi-Participant Learning-Environment Resource,” in *Networking and Complexifying the Science Classroom: Students Simulating and Making Sense of Complex Systems Using the Hubnet Networked Architecture*, U. Wilensky, (Chair) and S. Papert (Discussant), at the annual meeting of the American Educational Research Association, San Diego, CA.

Abrahamson, D., and Wilensky, U. (2005), “The Stratified Learning Zone: Examining Collaborative-Learning Design in Demographically-Diverse Mathematics Classrooms,” in *Equity and Diversity Studies in Mathematics Learning and Instruction*, D. Y. White (Chair) & E. H. Gutstein (Discussant), Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Eizenberg, M. M., and Zaslavsky, O. (2003), “Cooperative Problem Solving in Combinatorics: The Inter-Relations between Control Processes and Successful Solutions,” *The Journal of Mathematical Behavior*, 22, 389–403.

Fuson, K. C., and Abrahamson, D. (2005), “Understanding Ratio and Proportion as an Example of the Apprehending Zone and Conceptual-Phase Problem-Solving Models,” in *Handbook of Mathematical Cognition*, ed. J. Campbell, New York: Psychology Press, I. 213-234.

National Center for Education Statistics, National Assessment of Educational Progress (NAEP) (2004). *1996 National Performance Results*. Accessed March 5 and 23, 2004.
nces.ed.gov/nationsreportcard/ITMRLS/

National Council of Teachers of Mathematics Academy (2004). *Data Analysis and Probability*. Accessed March 4, 2004.

standards.mctm.org/document/chapter6/data.htm

Papert, S. (1991), "Situating Constructionism," in *Constructionism*, eds. I. Harel and S. Papert, Norwood, NJ: Ablex Publishing, I. 1-12.

Piaget, J., and Inhelder, B. (H. Weaver, trans.) (1969), *The Psychology of the Child*, NY: Basic Books.

Surowiecki, J. (2004), *The Wisdom of Crowds*, New York: Random House, Doubleday.

Wilensky, U. (1993), "Connected Mathematics—Building Concrete Relationships with Mathematical Knowledge." Unpublished doctoral dissertation, M.I.T., Cambridge, MA.

Wilensky, U. (1995), "Paradox, Programming and Learning Probability," *Journal of Mathematical Behavior*, 14, 231-280.

Wilensky, U. (1997), "What Is Normal Anyway?: Therapy for Epistemological Anxiety," *Educational Studies in Mathematics*, 33, 171-202.

Wilensky, U. (1999), "NetLogo." The Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.

Wilensky, U. (2001), "Modeling Nature's Emergent Phenomena with Multi-Agent Modeling Languages," in *Proceedings of Eurologo 2001*, Linz, Austria.

Wilensky, U., and Stroup, W. (1999a), "HubNet." The Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.

Wilensky, U., and Stroup, W. (1999b), "Participatory Simulations: Network-Based Design for Systems Learning in Classrooms," in *Conference on Computer-Supported Collaborative Learning*, Stanford University, Stanford, CA.

Dor Abrahamson
Graduate School of Education
University of California
Berkeley, CA 94720-1670
U.S.A.
dor@berkeley.edu

Ruth M. Janusz
Nichols Middle School
Evanston, IL 60202
U.S.A.
JanuszRM@aol.com

Uri Wilensky
Center for Connected Learning and Computer-Based Modeling

Northwestern University
Evanston, IL 60208-0001
U.S.A.
uri@northwestern.edu

[Volume 14 \(2006\)](#) | [Archive](#) | [Index](#) | [Data Archive](#) | [Information Service](#) | [Editorial Board](#) | [Guidelines for Authors](#) | [Guidelines for Data Contributors](#) | [Home Page](#) | [Contact JSE](#) | [ASA Publications](#)