SAMPLER: Collaborative Interactive Computer-Based Statistics Learning Environment[i]

Dor Abrahamson and Uri Wilensky
abrador@northwestern.edu uri@northwestern.edu
The Center for Connected Learning and Computer-Based Modeling

S.A.M.P.L.E.R.—Statistics As Multi-Participant Learning-Environment Resource—is a *participatory simulation* (Wilensky & Stroup, 1999a). In participatory simulations, a classroom of students collectively simulates a complex phenomenon that they are studying, with each student playing the role of a single agent or a set of agents in this phenomenon. For example, students may each "be" an atom in a molecule, a bird in a flock, or a sample in a distributed population. Technology-based participatory simulations are built in the NetLogo (Wilensky, 1999) cross-platform agent-based modeling-and-simulation environment and operate through the HubNet (Wilensky & Stroup, 1999b) architecture. Each student operates a NetLogo computer "client" connected through wireless hubs to the NetLogo "server" (in other HubNet activities, students operate calculators). This server "scoops up" student input, processes and displays this collective input, and sends messages back to the students' computers. The activity moderator, e.g., the teacher or student leader, controls the simulation parameters. So students embody agents in the virtual simulation they see projected onto a classroom screen. Also on this screen, monitors and plots display mathematizations of the simulation, so the class can explore relations between the model's initial conditions, student–agent rules of behavior, and collective outcomes. One example of a HubNet participatory simulation is "Disease," in which students "become infected" and then infect others in their virtual population. In "Gridlock," another example, each student controls a traffic light in a city grid. SAMPLER is a mathematics participatory simulation. It is part of the NetLogo ProbLab model-based curriculum under development (Abrahamson & Wilensky, 2004a, 2004b). SAMPLER classroom activities are designed to help students: (a) ground statistics in visual proportional judgments, stochastic-process intuitions, and additive and multiplicative action models; (b) reconcile micro–probabilistic (samples) and macro–gestalt (population) levels of reasoning; and (c) negotiate and integrate these personal resources through interpersonal discourse. SAMPLER statistical activities involve collecting data that informs estimations of population metrics under conditions of uncertainty and limited resources.

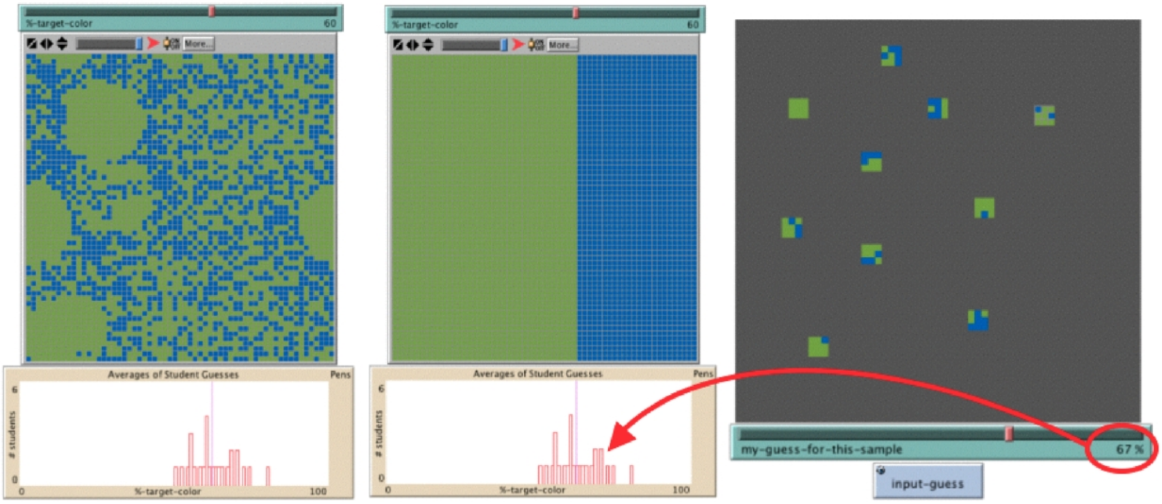


*Figure 1*. NetLogo S.A.M.P.L.E.R.: fragment of server interface with revealed population mosaic (scrambled, on the left, or 'organized,' in the middle) and histograms of students' guesses of the color density and mean of these guesses (note alignment with population green–blue contour). On right is a fragment of the student interface with personal samples from the population. The average color density in these ten 3-by-3 samples is 60/90 and the student has input 67%.

## Design of Classroom Activities

The population in SAMPLER is a large square "mosaic" of thousands of green and blue tiles (NetLogo "patches," see Figure 1, on left). The population property in question is the overall percentage of green or the greenness in the population. Alternatively, one might couch the issue in terms of the chance of a single hidden tile to be found as green. Yet, one might adopt a mid-level perspective and focus on the greenness of 9 cells, or 25 cells, etc.

Working in SAMPLER, students use their individual computers to sample from the shared population (see Figure 1, on right) and then each student determines and inputs a calculated guess of the overall greenness based on his/her samples. Student input is pooled and displayed as a histogram. If enough students are participating, then even if some of the guesses are outliers, the histogram's central tendencies will be quite accurate, stimulating discussion about the advantage of collaboration, pooling resources, and strength in diversity of opinions.

SAMPLER units begin with two preparatory stages that familiarize students with the learning environment. Initially, students observe and interpret a classroom projection of the population without the help of their own computers. The population is uncovered so all the green and blue tiles are visible (Figure 1, on left). Students initiate and practice strategies for gauging, measuring, and indexing the greenness of the population, each student guesses the greenness, and students suggest, evaluate, and debate methods for establishing a class guess (e.g., sharp-eyed student's guess, teacher's guess, the average). Once a class guess has been established, for instance by calculating the mean of all student guesses, the moderator "organizes" the population (see Figure 1, middle). The 'organize' function rearranges the mosaic tiles, putting all the green tiles on the left and all the blue tiles on the right. Students are encouraged to see how the organized population self-indexes its greenness along a left-to-right continuum. For instance, in the organized population in Figure 1 (middle image), the green extends 2/3 of the population's width. Next, the moderator creates a new population with a random, unknown greenness. The moderator demonstrates a sampling action by mouse-clicking on the server interface at locations the students select. Each click reveals a square sample around the cursor. Students discuss how to optimize sample parameters, such as location, size, and number of samples, so as to maximize the accuracy of their guess that is based on these samples. Finally, and during most of the unit, each student, clicking on their personal monitor, samples from a new hidden population (so students each see only their own samples). Students may choose to work alone, in pairs, or in groups. In between rounds, the teacher leads conversations about sampling strategies. Students receive a limited sampling "allowance" (how many tiles in the population they may reveal per round). Also, students receive initial points and then points are deducted for imprecision, 1 point for every 1% off mark. Students may discover that it is worthwhile to pool information and "gamble" on the group guess, but some students may wish not to do so in order to get ahead of the group. SAMPLER is designed to facilitate and mobilize classroom group dynamics: students are encouraged by the design to engage in interpersonal negotiation of facts and skills. Just as samples may be individually different yet must all be embraced in quantifying population metrics, so each student's voice may be unique yet equally important as all other classroom voices. In a sense, the mathematical machinery of statistics that is embedded in the design of the SAMPLER participatory simulation projects onto the classroom forum as a model of diversity, equity, and collaboration.

## Design Principles

We strive to build learning environments that foster students' grounding deep understanding of mathematics in accessible and engaging contexts (Confrey, 1993; Lesh & Doerr, 2002; Papert, 1980; Wilensky, 1997). Based on our previous research into student intuitions about probability (Wilensky, 1993, 1995, 1997), we designed SAMPLER so as to foster a community of users who: (a) experience a complementarity of probability and data analysis (Biehler, 1995); (b) distinguish and connect mathematical and empirical probability (Henry, 2001); (c) see a population as an aggregate of successive sampled events (keeping all the raw data rather than representations of the data); (d) experiment through re-sampling (Simon & Bruce, 1991; Konold, 1994); (e) express probability as base rate (Gigerenzer, 1998); (f) understand statistical patterns as emergent phenomena (Wilensky, 1997); (g) engage judgments of proportion (Spinillo & Bryant, 1991) and density (Gelman & Williams, 1998) as natural qualitative-cum-quantitative interpretations of visual information; (h) ground stochastic processes in spatial metaphors (Abrahamson & Wilensky, 2003); (i) analyze color density/intensity as a metric affording conceptual continuity between micro and macro levels of reasoning (from small sample to large sample to whole population); (j) re-invent mathematical strategies in response to challenging problems (diSessa, Hammer, Sherin, & Kolpakowski, 1991); (k) use fairly simple arithmetic to broaden the zone of classroom inclusion (Fuson et al., 2000); (l) build a cohesive, coherent, and fluent conceptual domain through mathematizing phenomena and, reciprocally, 'storyizing' mathematical representations (Fuson & Abrahamson, 2004); and (m) utilize collaborative group dynamics to connect to statistical constructs (Wilensky, 1993).

## Design-Research Methodology, Implementation, and Findings

We maintain that SAMPLER engages and carries students' probabilistic–statistical intuitions towards mature domain fluency. This section presents data to support this claim. We have worked with two focus groups and a pilot 6th-grade classroom, with students ranging in age (10 to 13), SES, ethnicity, and mathematics achievement. Each implementation informed modifications in the design towards the subsequent implementation. The students were told they would play mathematics learning games. The design-research team conducted the interventions according to a lesson plan that evolved into the plan laid out earlier in this paper. Throughout the intervention, students were encouraged to share and discuss their opinions. All interventions and pre- and post-interviews were video- and audio-taped for close analysis, and extensive parts were transcribed and debated in our research team. The following findings are based on microgenetic analysis of the data. We discuss only themes that recurred across interventions and many participants. The themes are mathematical, social, metaphorical, and combinations thereof.


*Figure 2*. A 6th-grade student interpreting the SAMPLER population of green and blue tiles. On the left and in the middle, the student is explaining a histogram that compiles his classmates' guesses of the population's greenness (the tall thin line between the histogram bars is the mean guess). On right, the student discusses how the scrambled population that is 50% green became "organized" (all green on the left, all blue on the right) to facilitate comparison between the histogram mean and the population greenness (note how the green–blue contour indexes the greenness).

### *Group Theory of Social–Statistical Interaction*

A surprising analogy emerged between student intuitions regarding the two axial populations inherent in SAMPLER: the mosaic as a collection of samples and the classroom as a collection of guessers. Students' anticipation of the validity of their collective guess interacted with these aggregates. On the one hand, students expected an increase in the variability of the distribution of guesses as a result of an increase either in the number of samples or of guessing students. On the other hand, they argued that whereas adding samples always increases accuracy, adding humans beyond some critical number of peers decreases guessing accuracy. So a design that taps group dynamics in mathematical practices must negotiate social sensitivities.

Group guessing modeled mutual liability in democratic societies: diligent students complained when taxed by the errors of classmates. Also, students discovered that in order to avoid a biased distribution of guesses, it is better to “guess first, discuss later” rather than “discuss first, guess later.” That is, first bidders—in particular charismatic and/or high-achieving students—may sway the group towards an anchor point, around which the group converges. Perhaps the most curious debate occurred after three students saw that whereas their histogrammed guesses were clustered, their fourth group-mate’s guess was “way off.” At first, they accused him of ruining the group guess. Lo, once we revealed the true population value, it turned out that all 4 students as a group had achieved a perfect hit. The outlying student took full credit for having tugged the group towards the hit. His three peers had mixed feelings: they were not sure how to judge him now, since he was still more “off” than each one of them individually.

*The Law of Large Social Numbers*

Students must take a leap of faith in order to surrender themselves to the power of aggregate guessing. In particular, students struggle with what is for them an apparent paradox: an increase in the number of guessers increases the cluster of guesses around the mean, yet it also increases the range of all guesses. Perhaps the law of large numbers is difficult to understand precisely because students have difficulty reconciling this apparent paradox. It is a challenge of our design to articulate, clash, and synthesize these would-be irreconcilable beliefs through simulation and dialogue. One interesting direction to pursue is that of the notion of compensation or balance, which students were groping at—for every guess that is *x* percentiles “way off” to the left there is another that is equally deviant to the right (see also Wilensky, 1997).

*Figure 3.* Students of diverse mathematical skill engage in sampling and calculating population metrics in the SAMPLER participatory simulation. From left: students (a) compute the mean of their samples; (b) lead the class by taking samples from the server; and (c) share a computer in taking their own samples from the population.

*Equity and Inclusion*

Because all students—the more and the less “number crunchers”—work and guess under shared uncertainty, alternative mathematical thinking is embraced as long as it is effective. For example, a 10-year old low-achieving student gained new social status through innovative pragmatic strategies, even though she could not quite explain the strategies in mathematical nomenclature.

*Sampling Strategies*

Students developed complex strategies to maximize the effectiveness of their sampling distribution over the mosaic. Students debated over optimal tradeoffs between few–large and many–small samples. Students were particularly concerned with strategies for addressing populations with non-random distribution of green and blue tiles (see Figure 1, on left). Those students who had argued that strategic distribution of samples over the mosaic is irrelevant to the accuracy of sampling re-evaluated their hypotheses once the entire population was revealed.

Some students who spontaneously applied a many–small sampling strategy, appeared comfortable in basing their approximation of the population’s greenness on a quasi-qualitative impression accumulating over repeated sampling (‘base rate’ strategy). For example, one 10-year old student rapidly re-sampled 3-by-3 arrays of tiles (“9-blocks”) that each vanished the moment a new sample was taken. While sampling, this student uttered the number of green tiles in each sample—“3, 0, 1, 2, 2, 1, 2, 3, 3, 1, 0, 2, 3, 2”—and then stated that about 2/9 of the population were green. This strategy was emulated and improved by another student who, working at the maximum speed the software allowed her, took numerous repeated samples of size 1 tile.

Students showed great flexibility and creativity in their sampling strategies. For instance, at a moment when 5 samples were revealed on the screen, each of 25 tiles, one student computed that each tile was weighted as .8% of the data because 100/125 = .8. When the total number of green tiles in the data was tallied at 64, his peer computed 64 * .8 and they inferred a 51.2% greenness in the population. These students were comfortable, for later samples, to work in a manner which, using the same numbers as above, would look as following: first count up the total tiles (125) and the total green tiles (64) and then divide 64/125, to get 51.2%. Yet another strategy, initiated by a 10-year old student, was to plan how many tiles she would reveal in total and then figure out how many green tiles would comprise 1% of this total, e.g., 3 tiles are 1% of 300 total sampled tiles. As she sampled, she eyeballed each sample, adding up 1 percent for every additional 3 green tiles that were revealed. Although she worked rapidly, nonchalantly, and not too accurately, her strategy proved highly effective, probably because her errors compensated one for the other.

*Agent–Aggregate Density Relations*
SAMPLER populations can be set as not bi-linear randomly distributed (see Figure 1, on left), creating clumps of green and/or blue tiles. Students were wary of hasty inferences in the absence of what they deemed as sufficient information. Their strategies and discourse revealed both an interpretation of samples as instances within a distribution and sensitivity to typicality of samples. Small- and medium-sized samples were not expected to be representative of the population, especially when initial samples revealed high variability in local density.

*Spatial Metaphors and Statistical Reasoning*
In order to contextualize students' discourse and rationalize the sampling procedure, we set the following task: "I am the CEO of an international shampoo company and I've assigned you, my statisticians, to determine the average number of hairs on peoples' heads around the world." Fortuitously, the mosaic resembled a map of the world, with the green clumps being continents in a blue sea. So context guided metaphor, and metaphor served as a vehicle of reasoning. One student said that if we concentrate our limited sampling resources in "South America" (green clump on lower left hemisphere of the mosaic-as-map), our inferences would not generalize to other continents. Another student localized the metaphor to sampling from different states in the U.S. We suggested a sampling strategy by which all samples should be taken from a remote town in the north of Finland. Students were not stumped. As one 10-year-old student objected, "What if they're all bald?"

Other interpretations students suggested for the mosaic were: (1) a maze; (2) "a green plain viewed from above and people wearing square blue hats"; (3) "a patchwork quilt"; (4) a representation of fashion styles over a whole year [sic]; (5) "the blue is small flocks of sheep and the green is hundreds and hundreds of walls"; (6) "the blue is flowers and the green is lots of forest plants that eat flies." These imaginative metaphors students construct for green and blue tiles suggests that: (a) the mosaic does not necessarily constrain students to a narrow conceptualization of the statistical construct "population"; and (b) any notion that students need explicit imagery so as to be engaged and making sense of a learning activity and that nondescript visual cues are too abstract is at the least questionable (see also Wilensky, 1991; Uttal and DeLoache, 1997).

*Living With Uncertainty*
Although students playfully guessed the greenness of *individual* samples, their attitude gradually shifted towards less commitment to samples and more to the population. That is, participants' intellectual investment and personal stakes centered on inferences concerning the *aggregate* and not on guesses concerning each successive bit of local data. For instance, samples that were "way off" caused surprise and updating of macro values but did not call for re-evaluation of strategy and did not harm students' esteem as mathematicians. Also, students learned to compile samples and to average classroom guesses so as to approximate the true population mean, which "you can never be certain of." This tenuousness of statistical truth informed students' seeing sampling as a

compromise between limited resources and desired accuracy. Also, students often referred to the time factor. Yes, they could count up thousands of tiles in a revealed mosaic to achieve 100% accuracy, but "that would take days." In that, time became as important a resource to consider in sampling strategies as were sample size, location, and number. Time interacts with these three spatial attributes because it dictates efficiency.

*Intuitive Statistics*

Students' spontaneously intuited mathematical constructs closely shadowed a host of real statistical concepts. Using simple quantitative vocabulary, students: (1) re-invented *margin of error* and *confidence interval*; (3) sensed that a mean is not always meaningful if the data is not *distributed normally*; (4) manifested and argued for sophisticated *sampling strategies* that address the gauged *distribution of target values in the population*; and (5) specifically expressed concern that samples from clumpy populations are problematic ("you should spread it out!"). Thus, probability and statistics can be seen as a high-precision enhancement of common-sense calculated estimation under conditions of uncertainty

## Conclusions and Future Work

SAMPLER engages students in activities wherein a shared object serves as a platform for articulating intuitions, learning professional vocabulary, testing hypotheses, and debating strategies of statistical inquiry. The inherently collaborative activities in SAMPLER, embodied primarily in students' interdependence for data and for estimates from these data—impelled students to scholastic argumentation that: (a) teased out individuals' intuitions; (b) afforded opportunities to engage in and refine mathematical terminology, representational forms, and conceptual tools; and (c) introduced and positioned 'distribution,' 'variability,' and complementary micro and macro perspectives in probability and statistics as cultural–-mathematical constructs. We conclude that <u>*advanced* statistical conceptual tools that are traditionally introduced as secondary school constructs already have their qualitative-cum-quantitative roots in elementary-school students' reasoning</u>. These roots can grow deeper and this reasoning can flourish, given appropriate learning environments.

We are planning to conduct further research on SAMPLER in middle school. In this context, SAMPLER will be implemented in conjunction with ProbLab (Abrahamson & Wilensky, 2004a, 2004b), a suite of curricular models that we are designing in collaboration with middle-school teachers. In ProbLab, students working with computer models in a collaborative learning environment: (a) reason about their own assumptions concerning randomness (see Abrahamson, Berland, Shapiro, Unterman, & Wilensky, 2004); (b) analyze and represent strings of random events (e.g., hit-miss-miss-hit-miss, etc.) using multiple plots (of *m/n*, attempts-until-hit runs, and samples); (c) manipulate and run experiments with mathematical objects that bridge between the familiar–concrete and the virtual–conceptual space of probability (e.g., computer-based dice and color boxes); (d) conduct combinatorial analyses of mathematical objects; and (e) use geometry to connect mathematical and empirical understanding of stochastic processes (Abrahamson & Wilensky, 2003). A key mathematical object in Problab is the square array of NetLogo "tiles" (e.g., the 3-by-3 "9-block," see samples in Figure 1). Students investigate and build, in both concrete and virtual media, the combinatorial sample space of all possible 9-blocks, where each of the 9 squares may be one of two colors (see Figure 4). The perceptuality of this histogrammed sample space may inform students' sense of sample distribution in working with the ProbLab models.

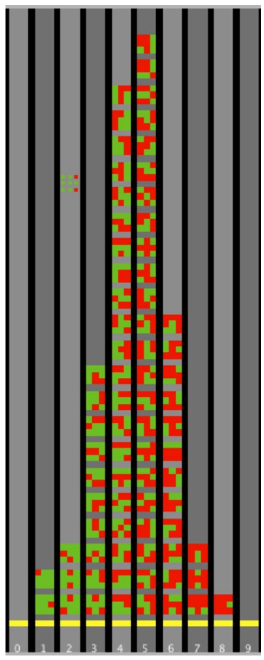

*Figure 4*. Fragment from NetLogo ProbLab Sample Stalagmite. The combinatorial sample space of 3-by-3 arrays grows probabilistically.

References


Abrahamson, D., Berland, M.W., Shapiro, R. B., Unterman, J. W., & Wilensky, U. (2004). Leveraging epistemological diversity through computer-based argumentation in the domain of probability. *Proceedings of The Sixth International Conference of the Learning Sciences*, Santa Monica, June 22 – 26, 2004. (Manuscript in press) http://ccl.northwestern.edu/papers/Abrahamson_et_al_ICLS04.pdf

Abrahamson, D. & Wilensky, U. (2003). The quest of the bell curve: A constructionist approach to learning statistics through designing computer-based probability experiments. *Proceedings of the Third Conference of the European Society for Research in Mathematics Education*, Bellaria, Italy, Feb. 28 – March 3, 2003. http://ccl.northwestern.edu/cm/papers/bellcurve/

Abrahamson, D. & Wilensky, U. (2004a). *ProbLab: Multi-agent interactive computer models for grounding probability in perceptual judgments of spatial proportions and in accessible mathematization*. (Manuscript under review).

Abrahamson, D. & Wilensky, U. (April, 2004b). S.A.M.P.L.E.R.: Statistics As Multi-Participant Learning-Environment Resource. In S. Papert (Discussant) *Networking and complexifying the science classroom: Students simulating and making sense of complex systems using the HubNet networked architecture*. The annual meeting of the American Educational Research Association, San Diego, CA, April 12 – 16, 2004.

Biehler, R. (1995). Probabilistic thinking, statistical reasoning, and the search of causes: Do we need a probabilistic revolution after we have taught data analysis? *Newsletter of the international study group for research on learning probability and statistics, 8(1).* Accessed Dec. 12, 2002: *http://seamonkey.ed.asu.edu/~behrens/teach/intstdgrp.probstat.jan95.html*.

Confrey, Jere (1993). The role of technology in reconceptualizing functions and algebra. In B. Pence (Ed.) *Proceedings of the 27th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education.* Asilomar, CA.

diSessa, A. A., Hammer, D., Sherin, B., & Kolpakowski, T. (1991). Inventing graphing: Meta-representational expertise in children. *Journal of Mathematical Behavior*, *10(2)*, 117 – 160.

Fuson, K. C. & Abrahamson, D. (2004). Understanding ratio and proportion as an example of the Apprehending Zone and Conceptual-Phase problem-solving models In J. Campbell (Ed.) *Handbook of mathematical cognition.* (Manuscript in press).

Fuson, K. C., De La Cruz, Y., Smith, S. B., Lo Cicero, A., Hudson, K., Ron, P., & Steeby, R. S. (2000). Blending the best of the twentieth century to achieve a mathematics equity pedagogy in the twenty-first century. In M. J. Burke (Ed.), *Learning mathematics for a new century* (pp. 197 – 212). Reston, VA: NCTM.

Gelman, R. & Williams, E. (1998). Enabling constraints for cognitive development and learning: Domain specificity and epigenesis. In D. Kuhn and R. Siegler, (Eds.). *Cognition, perception and language*. Vol. 2. Handbook of Child Psychology (Fifth Ed). (pp. 575 – 630). W. Damon, Editor-in-Chief. New York: John Wiley and Sons.

Gigerenzer, G. (1998). Ecological Intelligence: An adaptation for frequencies. In D. D. Cummins & C. Allen (Eds.), *The evolution of mind* (pp. 9 – 29). Oxford: Oxford University Press.

Henry, M. (Ed.). (2001). *Autour de la modelisation en probabilities*. Paris: Press Universitaire Franc-Comtoises.

Konold, C. (1994). Understanding probability and statistical inference through resampling. In L. Brunelli & G. Cicchitelli (Eds.), *Proceedings of the First Scientific Meeting of the International Association for Statistical Education* (pp. 199 – 211). Perugia, Italy: Università di Perugia.

Lesh, R. A. and Doerr, H. (Eds.) (2002), *Beyond constructivism: A models and modelling perspective on teaching, learning, and problem solving in mathematics education,* Mahwah, NJ: Lawrence Erlaum.

Papert, S. (1980). *Mindstorms*. NY: Basic Books.

Simon, J.L. & Bruce, P. (1991). Resampling: A tool for everyday statistical work. *Chance: New directions for statistics and computing, 4 (1),* 22 – 32.

Spinillo, A. G. & Bryant, P. (1991). Children's proportional judgments: The importance of "half". *Child Development*, *62*, 427 – 440.

Wilensky, U. (1991). Abstract meditations on the concrete and concrete implications for mathematics education. In I. Harel & S. Papert (Eds.), *Constructionism.* Norwood (pp. 193 – 204), NJ: Ablex Publishing Corporation.

Wilensky, U. (1993). *Connected mathematics—Building concrete relationships with mathematical knowledge.* Doctoral thesis, M.I.T.

Wilensky, U. (1995). Paradox, programming and learning probability. *Journal of Mathematical Behavior. 14(2),* 231 – 280.

Wilensky, U. (1997). What is normal anyway?: Therapy for epistemological anxiety. *Educational Studies in Mathematics 33(2)* 171 – 202.

Wilensky, U. (1999). *NetLogo*. Evanston, IL. Center for Connected Learning and Computer Based Modeling, Northwestern University. http://ccl.northwestern.edu/netlogo/

Wilensky, U. & Stroup, W. (1999a). Participatory simulations: Network-based design for systems learning in classrooms. *Proceedings of the Conference on Computer-Supported Collaborative Learning*, Stanford University

Wilensky, U. & Stroup, W. (1999b). *HubNet*. Evanston, IL. Center for Connected Learning and Computer-Based Modeling, Northwestern University. http://ccl.northwestern.edu/netlogo/hubnet.html

---

[i] SAMPLER and ProbLab models are available for free download at http://ccl.northwestern.edu . Special thanks to Matt Goto and Ethan Bakshy for their help in programming SAMPLER.