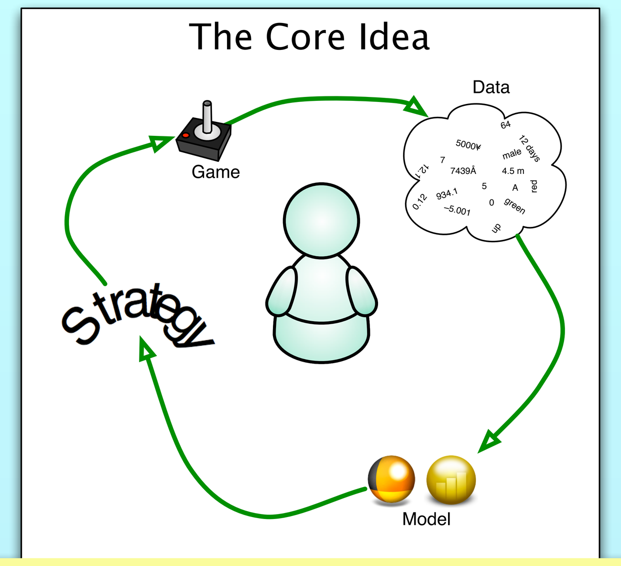
Data Games
Tools and Materials for Learning Data Modeling
Students improve their gaming by becoming data analysts of their own play.



Students playing computer games generate large quantities of rich, interesting, highly variable data that mostly evaporates into the ether when the game ends. What if in a classroom setting, data from games students played remained accessible to them for analysis? In software and curriculum materials developed by this project, data generated by students playing computer games form the raw material for mathematics classroom activities. Students play a short computer game, analyze the game data, conjecture improved strategies, and test their strategies in another round of the game.

Need: A Data-Literate Society
No solution to any significant problem facing us today—the global financial crisis, climate change, terrorism, poverty, and human rights abuses, to name a few—can be found without the efforts of people who can construct useful data models and without deep understanding of those models by people who adopt and apply them.

The research and development undertaken by this project address the need for more integration of data and chance with mathematics learning.

Goals

- Improve students' ability to understand and work with data, with special emphasis on large data sets, data visualization, time series, informal inference, and data structures beyond rows and columns.
- Enrich students' understanding of mathematics through learning experiences based on using data generated by playing computer games.
- Expand research in students' understanding of data and chance and in students' ability to learn mathematical content when it is embedded in data-rich contexts.

Foundations—Prior Work

Fathom—Dynamic data analysis software for secondary and college math and science classrooms.

TinkerPlots—Dynamic data exploration software for middle school mathematics classrooms.

TinkerPlots and Fathom have made powerful data visualization tools available to students throughout the US and in a growing number of other countries. Both development teams have been increasingly focused on the need to get data into the hands of students more quickly, for example through scraping data from a web page whose URL is dropped into a document. TinkerPlots is soon to release a new version that includes simulation capabilities.

Published by Key Curriculum Press

William Finzer, co-PI
bfinzer@kcptech.com
Vishakha Parvate, research and project manager
Kirk Swenson, software engineer
KCP Technologies
Emeryville, CA

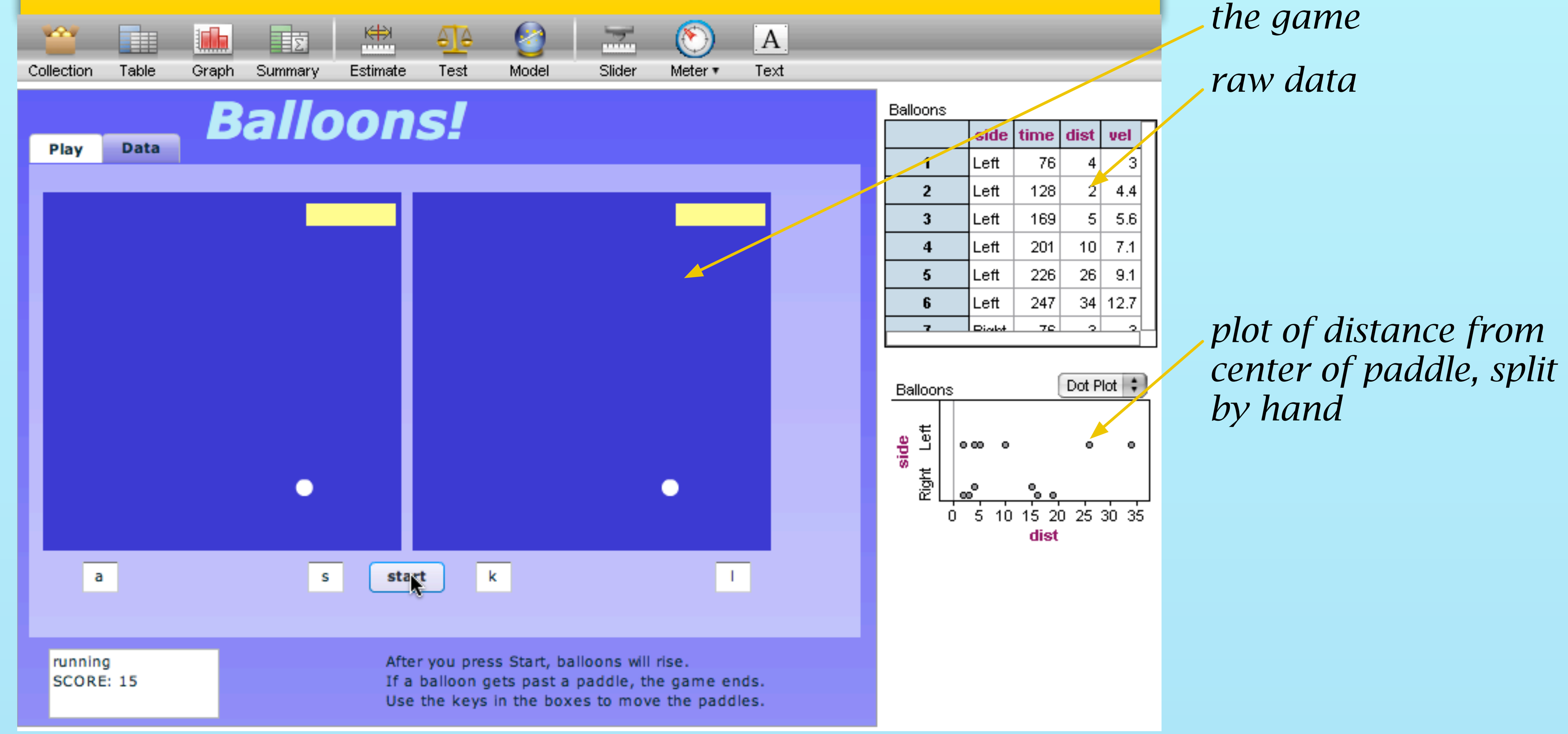
Cliff Konold, co-PI
konold@srri.umass.edu
Craig Miller, software engineer
UMass Amherst
Amherst, MA

Tim Erickson, activity developer
tim@eeps.com
Big Time Science
Oakland, CA

Balloons—An Example DataGame Activity
(a very early and rough prototype)

The Balloon Game Embedded in a Fathom Document
In this Pong-like game, the goal is to prevent the balloons from escaping by blocking them with paddles, one paddle for each hand. The balloons speed up as the game progresses.

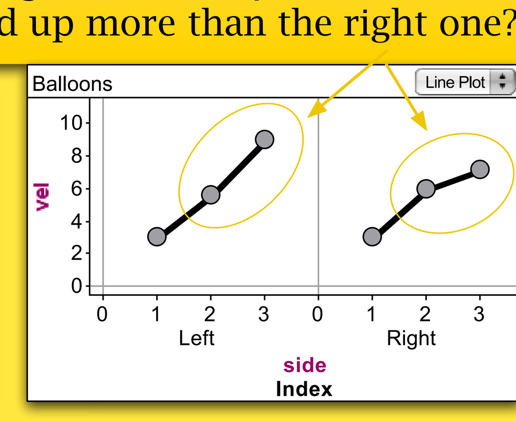
Data from the game flow into the Fathom (or TinkerPlots) document. The students' task is to build a data model of what is going on and use it to determine a strategy that will improve their game play.



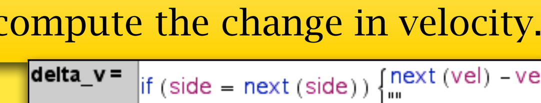
Construction of a Data Model

A First Strategy—Cover the Center
We decide to try to keep the paddles as close to the center as possible so the distance to move to block the balloon is as small as possible.

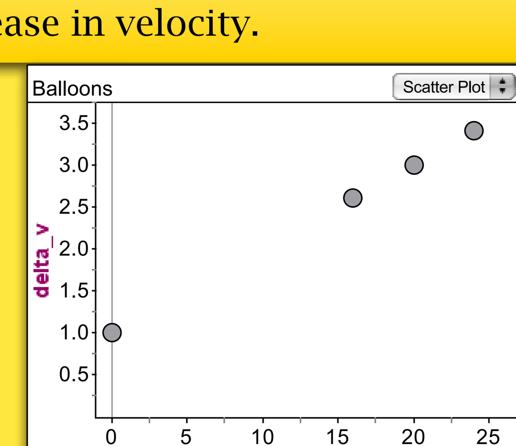
But in looking at the data, we notice that contrary to our expectation, the balloon velocities are not the same in the left and right hand. Why did the left balloon speed up more than the right one?



Investigation—ΔVelocity and Distance
What if the change in velocity has something to do with where on the paddle the balloon hits?
We compute the change in velocity.



Look for a Relationship
A graph shows that the closer to the center the balloon hits, the smaller the increase in velocity.



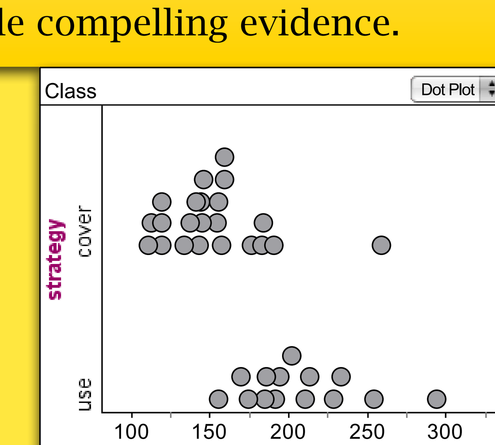
Strategy Revision

A Second Strategy—Use the Center
Since it's the high speeds that cause us to lose, and we can lower the speeds by blocking with the center of the paddle, our new strategy is to use the center as much as we can.

We try it, and it seems to work.

Can We Get More Evidence?
We enlist the rest of the class in an experiment. Each person will pick a strategy to play by coin toss. Scores will be collated automatically in a Fathom document accessible by anyone in the class.

The scores of the class as a whole provide compelling evidence.



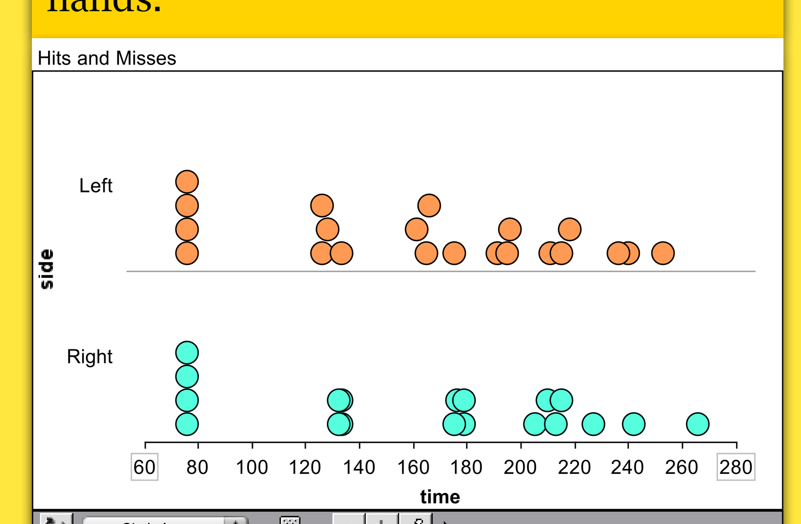
Activity Ingredients

A Computer Game
The game should

- Be quick to play
- Be fairly gender neutral
- Generate interesting data
- Have more than one useful strategy.
- Be more amenable to strategy than luck or repeated practice

Data Analysis
It is in analyzing the data generated by the game that students encounter the mathematical and/or statistical content of the activity. In this example activity possible content foci include rate of change, linear relationships, characterizing distributions, and experimental design.

Game Use at Multiple Levels
We expect that many games will prove useful in activities designed for both the middle and secondary school. For example, the balloon game could be used in a middle school activity in which the goal is to decide whether there is a difference between the right and left hands.



Research Questions

1. Students' conceptions of data that come from games: To what extent do students view the data as the result of a production process and does this conception have the same sort of affordances as repeated-measures contexts for interpreting data in terms of signal and noise?
2. Data structures: How do students view data, especially when they encounter data that do not fit into rows and columns (e.g. a binary tree of choices in a game)? What data structures are appropriate to introduce in middle school and, similarly, in secondary school? What type of user interface might mediate students' interactions with these data structures so that the structures are not barriers to data exploration and modeling?
3. Data visualization of large data sets: How do students' understandings, interpretations, and interactions with data change as a function of the size of the data set? What are the affordances of the various enhanced visualizations we develop in helping students explore and analyze data, especially large amounts of data?
4. Collaborative learning and data sharing: To what extent do the mechanisms the project builds for web-enabled collaboration and data sharing enhance classroom activities? What are the strengths and weaknesses of these mechanisms? What are appropriate directions for further development?

Mathematics Classroom Activities

Data Games will create and test ten activities for use in middle school mathematics and ten more for use in secondary mathematics classrooms. Each activity:

- lasts from one to five class periods;
- begins with brief play of a computer game embedded in Fathom or TinkerPlots;
- focuses on analysis of the data produced by playing the game with the goal of producing a data model of game play that can improve game strategy;
- builds mathematical and/or data analytic understanding through the data modeling (as opposed to game play);

Activities will be piloted in mathematics classrooms during years two and three of the project. After the project ends, activities will be made available for public use through Internet distribution in some form.

Software

Data Games will result in new versions of Fathom and TinkerPlots that:

- handle embedded Flash components and receive data and other information from them;
- allow data structures other than row-by-column, and provide interfaces for creation, manipulation, and visualization of these structures;
- provide new visualization techniques driven by data modeling of game data but suitable for many other kinds of data as well;
- make it possible for students to share game data and collaborate on data projects.

The software will be user-tested and piloted throughout the three years of the project. The plan is to have new commercial releases of Fathom and TinkerPlots about a year after the end of this project.