Kids' Survey Network: Teaching Data Literacy with Multiplayer Online Games
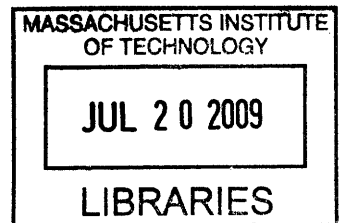
by

Christopher Wong

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Engineering in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

[June]
May 2009

ARCHIVES

Author . . ..
Department of Electrical Engineering and Computer Science
May 22, 2009

Certified by . . . .
Eric Klopfer
Associate Professor
Thesis Supervisor

Certified by . . . .
Scot Osterweil
Research Director
Thesis Supervisor

Accepted by . . . .
Arthur C. Smith
Professor of Electrical Engineering
Chairman, Department Committee on Graduate Students

Kids' Survey Network: Teaching Data Literacy with Multiplayer Online Games

By

Christopher Wong

Submitted to the Department of Electrical Engineering and Computer Science
on May 22, 2009, in partial fulfillment of the
requirements for the degree of
Master of Engineering in Electrical Engineering and Computer Science

Abstract

Current advancements in technology have permeated modern society, especially among the younger generations of computer users. Today's youth have grown up accustomed to video games and online social networking as standard elements of life. In contrast, many people of all ages today do not have a good grasp of data literacy. Many people simply do not understand statistics, charts, graphs and other survey data, which are all very commonly used in all aspects of everyday life. Kids' Survey Network is a large project that aims to focus the comfort and proficiency that middle school students have with web technology into an effective education tool to teach data literacy. This thesis project entails the creation of a suite of multiplayer learning games that reinforce the various other components of Kids' Survey Network.

Thesis Supervisor: Eric Klopfer
Title: Associate Professor

Thesis Supervisor: Scot Osterweil
Title: Research Director

3

Acknowledgments

Table of Contents

List of Figures

1. Introduction

Kids' Survey Network is a project designed to develop and study an apprenticeship network for informal math and science learning. An apprenticeship network is a web-based learning environment that encourages groups of users to become skilled in a certain area by engaging in project-based assignments. Kids' Survey Network's educational goal is to teach data literacy to middle school aged students through a web community that offers projects and games that create an opportunity to practice and advance through learning objectives. This thesis project consists of the design and implementation of the multiplayer games used for the larger Kids' Survey Network project. In this chapter, I will introduce the preliminary concepts and context of this project. Firstly, I will argue for the need to teach data literacy, through explaining what data literacy entails and why it is a necessary concept for all students to understand. Next, I will describe a current concept known as "Web 2.0" and explain how Kids' Survey Network fuses this idea with its goal of teaching data literacy. Subsequently, I will show how computer games can be used to further education and what MIT's The Education Arcade's objectives are in this field. I will quickly define usability and explain its importance in this project. Lastly, I will give a summary of the rest of this thesis.

1.1 Why Teach Data Literacy?

"There are three kinds of lies: lies, damned lies, and statistics." – Benjamin Disraeli

The above quote, popularized by Mark Twain, encapsulates the distorted use of data and statistics in our world today. Statistics are used in almost every sector of industry and government, and in addition, our news media is saturated with statistics about important current events. People use data every day in order to form opinions about government policies, health practices, neighborhoods to live in, and various other issues. However, statistics are often misused and misunderstood. People mix up causation and correlation or are unaware of the context or statistical significance of the data that they are presented with. This problem is exasperated by the fact that statistics are often very persuasive to many people that may not fully understand the meaning of the data.

Data literacy is the ability to comprehend and utilize data to correctly and advantageously draw inferences in solving a problem or forming an opinion. Topics in data literacy include sampling, measurement, graphing, statistics, and research design. In the standard mathematics track in American K-12 college preparatory education, the focus is heavily skewed towards algebra, geometry, and calculus. Most Americans' education on data analysis only reaches basic concepts like mean and median.

For most college bound students, especially those not majoring in mathematics or engineering, the ability to critically analyze data is a more useful tool to conduct research and form arguments than the mathematics of algebra or calculus. Even for many science majors, mathematics is most commonly used when evaluating the results of experiments. The ability to conduct survey research is an incredibly useful tool in the fields of government, science, and commerce [1].

It is critical that young students be given the opportunities to learn how to ask and answer questions about a set of data, to understand the legitimacy of the values of statistics, and to use data to find answers to a particular problem. With practice utilizing data, starting from a young age, students will be much better equipped for success in their academic careers, and they will also develop the ability to make more well informed decisions in their daily adult lives, both professionally and personally.

1.2 What is Web 2.0?

Web 2.0 is a concept that epitomizes a trend of Internet use in recent years. The term Web 2.0 refers to the perceived second generation of web development that focuses on communication and information sharing among peers. Whereas Internet media used to be created primarily by corporations and institutions and conveyed to the common user, now Web 2.0 media is driven and created by any Internet user for any other user. The main idea is that the Internet is a different media from traditional outlets such as television because it is focused on creating an expansive online community where users communicate information, make transactions, discuss issues, and socially network with each other.

Web 2.0 has taken a huge foothold on Internet users as a whole, but especially on the youngest generation. Websites such as Youtube, Myspace, Facebook, and Twitter are incredibly popular among teenagers for because these sites enable them to communicate and be socially active via the Internet [2]. Social networking is a huge part of Web 2.0, and these online communities have become a huge part of daily life for hundreds of millions of Internet users.

Kids' Survey Network is an attempt to create an online learning social networking community. Students using Kids' Survey Network will become part of a web-based community of learning. The web-based infrastructure allows a large community with a variety of projects to work together, share ideas, and interact with mentors. Each user will participate in other users' projects and users of different levels of expertise will assist each other through a group learning environment. In addition, there will be a suite of multiplayer learning games that introduce and consolidate skills necessary to undertake survey projects. These multiplayer games, which are the element designated to MIT's The Education Arcade and this thesis project, help build teamwork and camaraderie to both provide a powerful way to convey and reinforce core concepts and to underpin the online community that is Kids' Survey Network.

1.3 Education Through Technology and Games

It is not difficult to see that video games have become a huge part of the daily lives of today's youth. In the month of February 2009, during a time of global recession, video game sales actually increased to $1.47 billion [3]. This cultural fascination with video games has been very controversial among both educators and politicians. Some worry that the intrusion of video games into the lives of children will negatively affect their responses towards traditional instruction and that video games instill poor values while decreasing attention spans. However, video games do not always oppose the goals of education. Instead, game-based education can capture the attention of students and convey material in a way that no other form of education can.

In fact, games such as the *Civilization* series or the *SimCity* series have been heavily praised for integrating useful academic knowledge into the activities of the game. For example, *Civilization* teaches the basics of many different forms of government, the effects of technological advances on society, and the importance of diplomacy. *Age of Empires* recounts the stories of many true historical figures and events in its campaigns. The *SimCity* series teaches the basics of urban planning and also explores the role of local government. These existing games are great examples of the educational value of games and show the potential for greater educational throughput with the use of video games.

While there have been many commercial games that incorporate elements of education within the prioritized goal of entertainment, some educators feel that the most promising developments are in games explicitly developed to support learning. Education through computer games has been attempted for decades, since the advent of the home computer. Computer games provide interactivity, motivation, and organization which are often much harder to achieve using traditional education methods. With games developed for education, there can be closer matches between the desired learning outcomes, the computer and human resources available, the educational game play, and also the potential supplementary learning experiences [4]. With a better understanding of how learning happens during game play and realizing how game play can support other educational outlets, games can have a much larger impact on the education of their players. The Education Arcade, the MIT Comparative Media Studies group that I am working with to create games for Kids' Survey Network, strongly identifies with this goal. The Education Arcade is concerned with games that promote learning through authentic and engaging

play. The games designed by The Education Arcade demonstrate the huge potential of video games with education and also dynamically address the educational needs of the players. With these ideas in mind, the games designed for Kids' Survey Network have clear objectives to teach core ideas of survey research and statistics (such as categorization, hypotheses, and graphs), in a fun and engaging way.

1.4 Usability

Usability is the ease of learning and using a product and the overall satisfaction of the entire experience. Usability can be measured in 5 qualities: learnability, memorability, efficiency, errors and satisfaction [5]. Learnability is the ability for the user to complete the tasks of the product during the first time they encounter the user interface. Memorability is the capacity of the user to repeat the tasks in succeeding interactions with the product. Learnability and memorability are coupled in this way as they both deal with having users know how to use an interface. Efficiency is the speed in which users can finish tasks when they know what they need to do. Errors are mistakes that a user make while using the interface, and usability is measured in terms of the frequency and recovery of errors made by users. Satisfaction is the overall enjoyment a user feels when using the product.

Usability is a very important concept to analyze the multiplayer games of this project. Because the games are geared towards middle school aged users, learnability and memorability are extremely important, because children are usually not exposed to complicated user interfaces or complex programs. Middle school aged users are also less likely to fully read and comprehend instructions,

and may lose interest in a game if they do not know how to play it. Children also may be more likely to have errors as they are not expert computer users, and they may become easily frustrated with the errors if recovery is not easy. Lastly, because the product is a game, satisfaction and efficiency for the time based games are extremely important. The whole point of educating through games is to make learning a pleasurable experience. These topics of usability will be used to describe features of the design of the games made for Kids' Survey Network.

1.5 Thesis Summary

Chapter 2 provides background information specific to this project. I will introduce TERC, the nonprofit organization behind Kids' Survey Network, and also discuss previous work related to this project. I will also introduce the technologies used to implement the games in the project, Adobe Flex and Smart Fox Server. Chapter 3 presents more details about Kids' Survey Network, including the mission, audience, desired outcomes, and entire project plan. In Chapter 4, I will describe the game *Categorize This* in full detail, including the design process and the implementation created. Chapter 5 explains the same features of the game *Conjecture Graph Spoons*. In Chapter 6, I will provide results of user testing of the games and in Chapter 7 I will discuss the future work that will be completed for this project. The contributions of this thesis will be summarized and concluded in Chapter 8.

2. Background

The first two sections of this chapter give background information about the educational organization behind this project, TERC, and two projects TERC has done that are the predecessors of the Kids' Survey Network project. The following sections provide an introduction to the technologies used in this project, Adobe Flex and Smart Fox Server.

2.1 TERC

TERC is a research and development nonprofit organization based in Cambridge, MA. TERC's mission is to expand and improve math and science education so that creative, rigorous, and reflective inquiry are a part of the lives of students from all communities. TERC works towards their mission by collaborating with the education community, developing educational practices and products, conducting research in classrooms and other learning environments, creating learning opportunities for all students, and applying technology to meet educational needs [6].

For over 40 years, TERC has been working in a large variety of projects geared towards improving math and science education. Every year, TERC's programs reach 3.5 million students in the United States and internationally, many in underprivileged communities. Kids' Survey Network is just one of many software products that TERC has created towards math and science education. Two of its successful products are described in the following section.

2.2 Previous Work

Two previous pieces of software designed by TERC that are considered predecessors of Kids' Survey Network are *Zoombinis* and *InspireData.* These two products serve as this basis of Kids' Survey Network, as they are educational mathematics software developed by two lead designers of Kids' Survey Network, Scot Osterweil and Chris Hancock, and thus have shaped the design process of this project.

2.2.1 Zoombinis
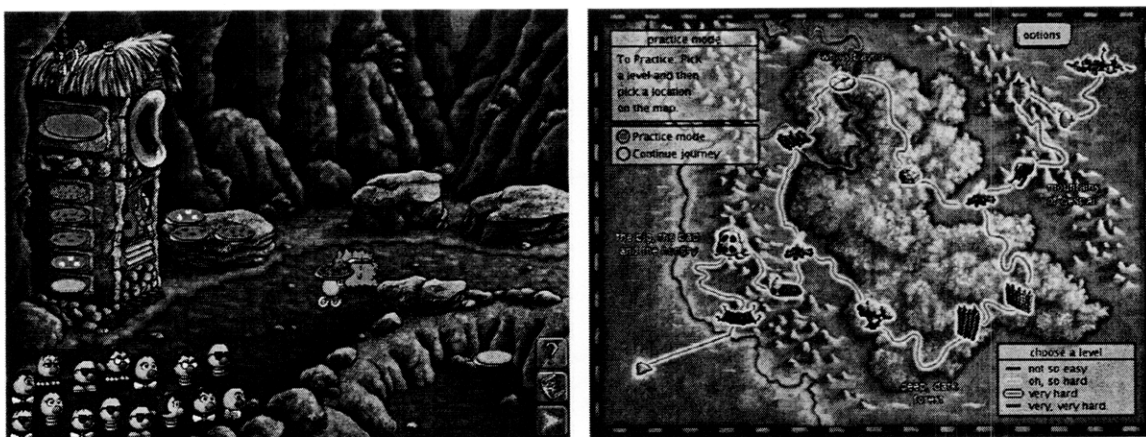
Figure 1: Screenshots from Zoombinis Logical Journey.

Zoombinis Logical Journey, along with its two sequels, is a critically successful game that has made a widespread impact in the educational software market. Over a million copies of *Zoombinis* have been sold since being released by Broderbund Software in 1996 and the game has won numerous awards. *Zoombinis* is structured as a series of puzzle games that consist of challenging

problems based on logic, permutations, and other discrete mathematical concepts. Key design characteristics that led to the success of *Zoombinis* include: coherence of content and story, depth of content, replayability, soft barriers, lack of time pressure, and a strong story [7]. The lack of time constraint enabled cooperation between many players, thus pioneering the way for group-based advancement in Kids' Survey Network. The depth of content and replayability concepts have also made their way into the multiplayer games for Kids' Survey Network in the form of multiple levels and dynamic game play from the multiplayer setting.

2.2.2 InspireData
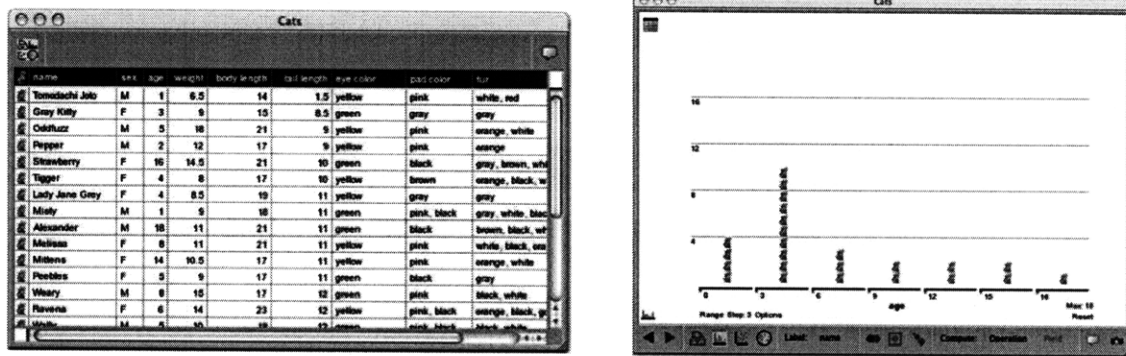
Figure 2: Screenshots from *InspireData*.

InspireData, while not a game, is a TERC-developed piece of software that is geared for building data literacy and data analytical skills in the classroom. *InspireData* is a very powerful visual learning technique that presents data graphically and interactively. Students use *InspireData* to examine and analyze data in an assortment of different charts and graphs. The software is highly

dynamic in the ability to change variables, ranges, or plot types to fully understand the best way to examine a particular set of data. Released recently by Inspiration Software, *InspireData* has also won numerous awards. The excellent graphical representation of data and the ease of use and interactivity are elements that Kids' Survey Network would like to inherit from this predecessor.

2.3 Adobe Flex

The games of Kids' Survey Network will be programmed using Adobe Flex, an emerging technology used by hundreds of organizations including MTV, Boeing, eBay, and the U.S. Department of Defense [8]. Adobe Flex is a technology used as an alternative to create SWF files (also known as Flash files). Flash games are visually engaging, animated, interactive games found very commonly on the Internet. Flash has become incredibly popular and is found on many websites, creating media-rich Internet applications. Most web browsers, and also other standalone desktop players, support the SWF file format. Having grown in popularity since its first release in 1996, Adobe claims that today, 98 percent of US web browsers support Flash.

While Adobe Flash (formerly Macromedia Flash) is the original technology used to create Flash animations and games, Adobe Flex is a more recently released development kit created by Adobe as an alternative to Adobe Flash. The Adobe Flash platform is designed as an animation metaphor geared towards designers, which creates a programming model that is unfamiliar to most software developers. Adobe Flex is Adobe's answer to the current demands

for rich Internet applications. Rich Internet applications are websites where the site actively responds and computes logic to create a streamlined flow of application use and deep interactivity between the website and the user [9]. While maintaining the ability to heavily incorporate graphics and animation, Flex creates an object oriented programming model (using the object oriented scripting language, ActionScript), with an excellent framework for designing user interfaces. In addition, Flex comes with a very comprehensive library for many useful user interactive components, including good handling of drag and drop, user input components, and moving and resizing animations. Flex puts the power and flexibility of a desktop application into an online site. In conclusion, Adobe Flash is a vastly popular, incredibly useful tool in making games as it creates great ease in integrating and controlling animations and other media.

2.4 SmartFoxServer

To handle the communications of the multiplayer games of Kids' Survey Network, we decided to use SmartFoxServer to handle the connections and communications between the multiple computers playing the game at once. SmartFoxServer is a mutli-platform socket server designed to integrate with Adobe Flash, allowing users to easily develop multiplayer applications and games. SmartFoxServer can handle thousands of high performance clients on a single server machine, which is the projected use of Kids' Survey Network. SmartFoxServer comes with an application programming interface in ActionScript that easily allows Adobe Flash or Flex programmers connect, communicate, and access data. The SmartFoxServer interface comes with both

user variables and room variables, and also allows for event propagation when updates are made to the multiplayer game state. SmartFoxServer is an incredibly useful server that makes programming multiplayer games in Flex much easier and allows for very deep interactions between the players of the Kids' Survey Network games.

3. Kids' Survey Network

In this chapter, I will present an overview of Kids' Survey Network, including the mission, the audience and projected outcomes, the components of the proposed apprenticeship network, and the project plan for the future. Lastly, I will go into detail about the specifics of the MIT involvement in the project, the multiplayer games.

3.1 The Mission of Kids' Survey Network

Kids' Survey Network is a project that works to combine its two goals of improving education levels and understand of survey research and also investigating and demonstrating the effectiveness of enjoyable, out-of-school educational contact.

The educational mission of Kids' Survey Network is to promote the rational, responsible, and successful use of data among students such that they will retain the skills necessary to employ data as citizens and workers. Surveys studies have become a vital tool for workers in government, business, and science. Many topics within survey research, such as sampling, measurement, graphing, statistics, and research design apply to many other useful areas, such as mathematics. These concepts and skills, sometimes called data literacy, can be acquired with the educated use of survey research. School-based education on this subject has lagged due to challenges with teacher training, curricular integration, technology integration, and alignment of testing. Kids' Survey

Network strives to make up for these difficulties by encouraging students and offering opportunities for them to actively apply the aforementioned data literacy concepts among their peers.

The strategic mission of Kids' Survey Network is to research how significant time spent in out-of school sessions can lead to learning for science, technology, engineering, and mathematics. Parents today are pressured to make their children spend extended time learning outside of school through extra programs, camps, and other organizations. Kids' Survey Network would like to discover how well enjoyable games can also double as a rigorous science and math related education platform.

To achieve these two intertwined missions, Kids' Survey Network will be an online learning environment that allows kids to use surveys to answer their own questions about the experiences and opinions of kids both across the country and in their local communities. Kids' Survey Network provides educational opportunities that local education programs could not produce on their own and allow the documentation of learning outcomes through children's participation. Associated research studies will be done to provide insight and evidence of the capacity for informal learning of science, technology, engineering and mathematics subjects.

3.2 Targeted Audience and Anticipated Outcomes

For the entire Kids' Survey Network project, the targeted audience is people aged 11-14 nationwide in out-of-school learning programs. An estimated 2 million people in the specified age group participate in these out-of-school learning programs, and during the initial phase of the Kids' Survey Network

project, 1,400 young people will participate via these groups. Young people aged 11-14 generally have very low data literacy but have shown high interest and comfort with computer games and online communities. These pre-teens and teens are motivated by community service and interest in their peer group, and many online instant polls are popular among them.

The anticipated outcome is to see a measurable improvement in the students' data literacy; including the ability to understand and perform graphing, descriptive statistics, and data-based reasoning. A second predicted outcome is to see the participants' grown appreciation and interest in the power of survey research and data-based information processing.

3.3 The Components of the Apprenticeship Network Model

Kids' Survey Network chooses to apply the apprenticeship network model to engage young learners in voluntary project-based learning. In the model, an apprenticeship network is an informal, web-based learning environment where groups of students learn while participating in a project-based endeavor. The project-based aspect of the learning model gives participants an overall intrinsic reward and incentive for their hard work. It also demonstrates that the skills they are learning are meaningful, reliable and respected. The group-based aspect of the model gives structure, companionship, and support to the individual.

There are four components used to organize and facilitate this practice: a web based community of practice, a common set of tools, a suite of learning games and tutorials, and structures for tiered, team-based advancement. The web-based community of practice allows various projects to exist between

different participants and brings them together to exchange ideas and interact with each other's projects. The common set of tools provides a standard technical language for the users to communicate ideas with each other. These tools enable users to create and manage surveys for manipulating and analyzing data as well as for producing presentations. To complement the survey tools, a suite of learning games and tutorials will provide a fun and easy way for the participants to introduce and strengthen skills and concepts needed in the survey projects. The tiered, team-based advancement means that the activities are not ordered in strict succession but integrated into new levels with different requirements and skill-developing trials that require teamwork and critical thinking.

3.4 Project Plan and Sustainability

Kids' Survey Network will follow a three year timeline to develop and roll out additions to the website until it is fully functional. The project will start in two design study sites to a national online network of 18 after school program centers with thousands of users. Every component of the site will be created and improved in an iterative pattern of design and development. The first 2 years will be focused on planning and development, and in the third year a summative evaluation will be done on the entire Kids' Survey Network project. After this 3 year development period, Kids' Survey Network will reach a funding period and reach 1,140 students in out-of-school programs.

Kids' Survey Network will become self-sustaining when its audience reaches 30,000 students. At this point, small fees may be paid by providers to sustain the project's core operations. The project will require heavy cooperation

between TERC, providers of after school services, contributors of out-of-school programs and research, and professional and academic organizations.

3.5 The Education Arcade's Role

The focus of my project is the multiplayer games described above as one of the four components to the Kids' Survey Network apprenticeship model. These online learning games offer a structured and motivating way to study and reflect concepts important to the rest of the project. The mini games will address key data literacy topics, and embeds the learning goal in a fun and narrative manner. The multiplayer aspect of the games will encourage teamwork and camaraderie, a huge goal towards guiding the groups of students in team-based advancement. The games will also bring an unexpected excitement to the important concepts that need to be imparted to the participants.

4. Categorize This!

Categorize This is the first multiplayer game we made for Kids' Survey Network. The design staff of The Education Arcade, led by Caitlin Feeley and Scot Osterweil, designed the game. The implementation was done by Colin Greenhill, Sam Potasznik, and myself.

4.1 Concept

Categorize This is designed to teach students how to define good data categories that are understood by other people. In this game, the players are all playing together on the same team, trying to score the highest they possibly can. The game involves the team members choosing 4 or 5 categories after seeing a sample set of images (from a much larger set of images). Then, the players individually sort the large set of images into the categories they thought up, with the goal of having the highest possible level of agreement between the different players. The varying levels of difficulty differ in the amount of time given to the players to choose categories and categorize, the total number of pictures, and the complexity of the set of images.

There are multiple learning objectives that this game aims to teach its players. The game teaches the importance of choosing meaningful categories that are mutually exclusive and that encompass the entire set of data. The idea of specificity is important such that categories shouldn't be too broad or too specific

so that the images are split into a good balance of categories. The game also stresses an idea called "inter-rater reliability," the level of how well the categories are intuitively understood by all the players, and transitively, understood by people in general. Lastly, the game requires players to predict characteristics of an entire set from a small sample set, and to deal with outliers that may not fit the simple categories that the players have chosen.

There are 3 main stages to the game: the category select stage, the object sorting stage, and the results view stage. After the initial log-in and room selection, a player is taken to the category selection stage. Each player is shown a small sample of images from the set of objects they will be asked to categorize. The players will discuss what categories to make, and the team leader (a role that switches with every game the team plays) enters up to 5 categories. A 4-minute timer counts down to provide a time constraint for this discussion. When all the players hit a "Ready" button to indicate their consent, the game moves on to the next stage.

The second stage of the game is the object sorting stage. At this point in the game, the chat functionality is disabled to keep the sorting of the images independent of player interaction. The players are given the same images in sets of 8; the order of the images among the sets of 8 is shuffled for the common situation where players are playing in the same vicinity. Each player is presented with one image at a time, which they drag into a column designated as the container for the images that fit into a particular category. The columns become a histogram, making a connection between a single point of data and a display of distribution. The game displays an agreement meter that indicates

higher scores if the players categorize images equivalently, and indicates lower scores if the players disagree.

The third and final stage of the game is the results screen. The players are congratulated for successfully completing the game, with a more enthusiastic and fun fanfare for more successful teams. The players are then shown a chart-based display revealing how each player categorized each image.

The seemingly simple task of categorization is in reality quite complex, even to adult researchers. It is especially hard for kids in the middle school range, as they often prefer to include as much information as they can. In addition, the concept of data distribution and outliers is a very advanced topic for students of the targeted age. To make these topics more understandable, we have chosen very common items to the participants (shoes, vehicles, pets, food) that make the categorization idea a lot more intuitive. These familiar images allow the players to discuss real issues of data categorization, make good predictions about the data set, and also understand the notion of outliers. The histogram formed by the images in the category boxes also builds the idea of a graphic representation of data, as the broad categories are filled with many images and the narrow ones are empty. The entire game streamlines all these learning concepts into a fast paced, fun, and intuitive game.

4.2 Front End Design

The design of the look and feel of the game was done with prototypes that expressed how the designers wanted the game to look like. The finished product did not stray far from the following prototypes.
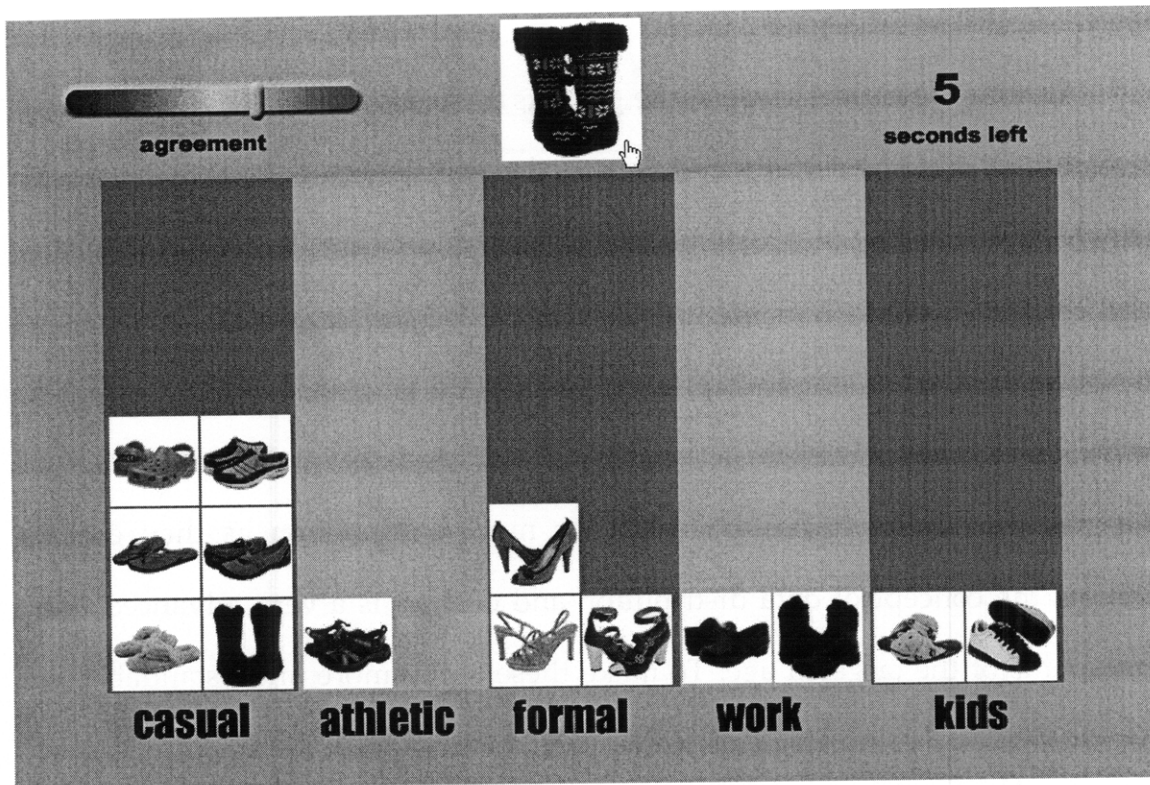
Figure 3: The prototype for the categorizing screen.

The categorizing screen was designed for both good usability and a good representation of the distribution of images into the categories. The prototype image shows the game in progress with a set of shoe images, and a histogram of shoe images distributed in good categories. The agreement meter uses the colors green and red for good and bad respectively. This familiar color paradigm allows for easy learnability, as the players can intuitively understand what the agreement meter shows. Also, the positive or negative feedback either reinforces or calls into question the behavior of the player. The image is to be dragged and dropped from its original position into one of the category bins. While drag and drop can sometimes have difficult learnability, as it is hard to convey that the

image is draggable, the tradeoffs are that the game has very high memorability, efficiency, and satisfaction. Dragging and dropping is by far more efficient than any other form of dropping the images into the categories, as it only requires one click and one drag. The images are snapped into their correct place after the release of the mouse button anywhere in the rectangle for a specific category, leading to high satisfaction in the drag and drop activity.
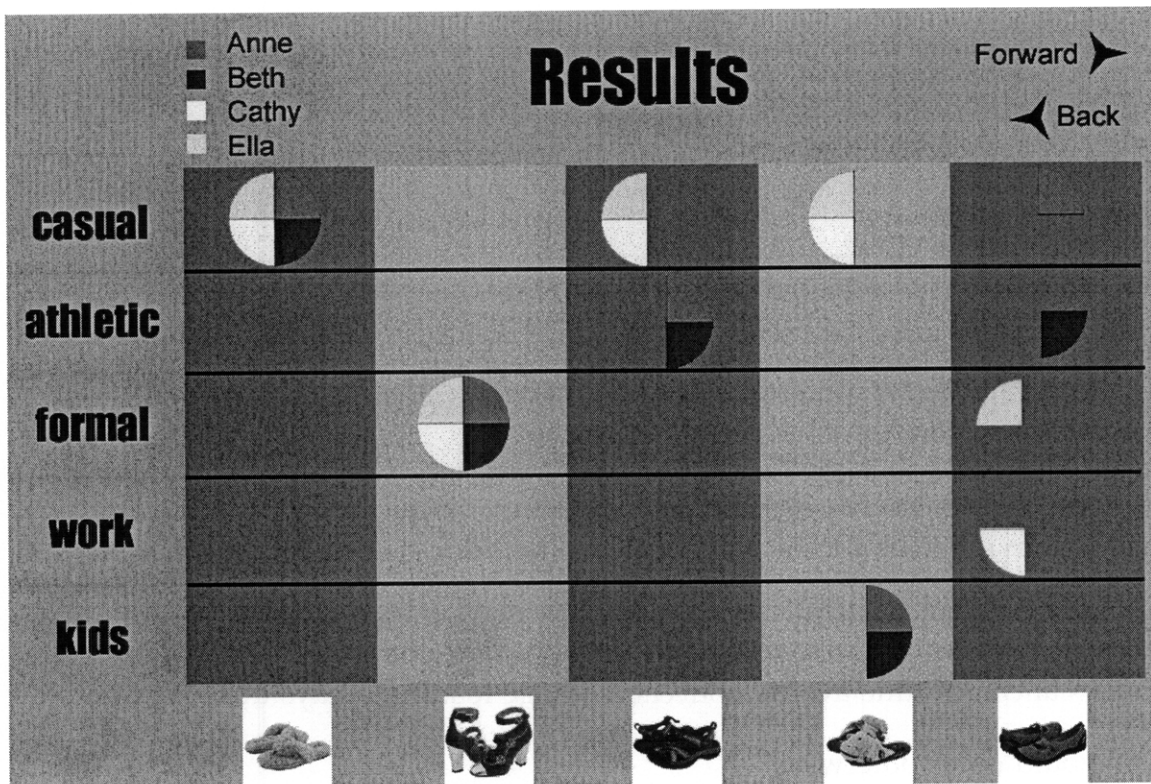

Figure 4: The prototype for the results screen.

Many options were considered for the manner in which to display the results, and the final implementation differed slightly from this late stage prototype. In this prototype, a pie chart was used for each image to show how much the team was in the agreement over the categorization of an image. Although the final implementation was slightly altered, it kept the main theme of

having a chart with some representation to show how well the players agreed for each image.

4.3 Back End Design

The software engineering design of *Categorize This* really was based around the front end user interfaces, and most of the code revolved around the user interface. In terms of the back end, the behind-the-scenes logic dealt with loading the deck, configuring shared variables, and sending the categorization data to compute the agreement and show the results.

The image sets are loaded into the game as XML files. We made this decision so that we can easily switch or alter the image sets (which are also the equivalent of game difficulty levels) with either minimal or nonexistent alterations to the source code. In addition, even individuals without any programming experience can create and modify the XML files. The XML file is loaded immediately after a player selects the desired level, and at that point in the game, the images are preloaded into each client's memory. Loading the images ahead of time avoids noticeable image loading delays during game play that would occur otherwise.

A few room variables were put into place, mainly to update whenever players were ready to move onto the next stage in the game. The biggest back end computational challenge was the aggregation of results, since every player needed to know each other player's exact categorization in order to fulfill the design of the results screen. After each player has categorized an entire set of 8 images, a mapping between images and categories is sent to the "host" player

who compiles all of the data. This host computer then sends the final results (as a total collection of each image and how it was categorized) to all the other players after all of the categorization has been completed and sent in.
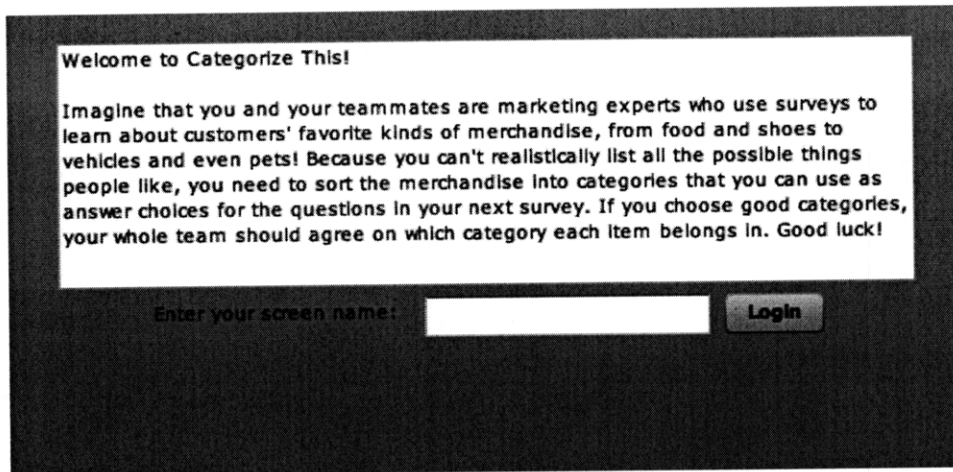
4.4 Implementation

Welcome to Categorize This!

Imagine that you and your teammates are marketing experts who use surveys to learn about customers' favorite kinds of merchandise, from food and shoes to vehicles and even pets! Because you can't realistically list all the possible things people like, you need to sort the merchandise into categories that you can use as answer choices for the questions in your next survey. If you choose good categories, your whole team should agree on which category each item belongs in. Good luck!

Enter your screen name: [   ] **Login**

Figure 5: The introduction screen to Categorize This!

The game of *Categorize This* is currently in an early "beta" stage of the software release cycle, meaning that the software has been released for controlled user testing among middle school aged students, and that all usability features have been implemented. None of the final artwork or any visual design has been put into consideration yet, as professional graphic designers will be brought in to help with the final polish of the game. In addition, as the games are built to be part of the larger Kids' Survey Network website that is unimplemented at this time, the login and room finding have been simplified for ease of use with testing. Figure 5 shows the simplified login and introduction screen currently available in this beta version.
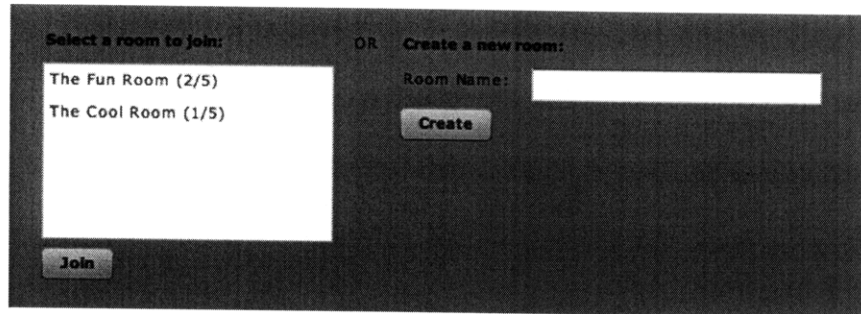
Figure 6: The room joining functionality.

The ability to support multiple games ("rooms") at once is supported by the current version of our game. As the room creation and assignment will be handled by the larger Kids' Survey Network website in the final release of the game, the current version has a simple room selection shown in Figure 6. After at least 3 players have joined the room (out of a maximum capacity of 5), the host player, which is initially the player who created the room, can choose to begin the game, where he or she remains the host player for the first round of play. When the game begins, all participants are brought to the category select screen. To provide adequate information to ensure user readiness in a time based situation, a warning message with a countdown is displayed before the transition to the screen. This display is shown in Figure 7.
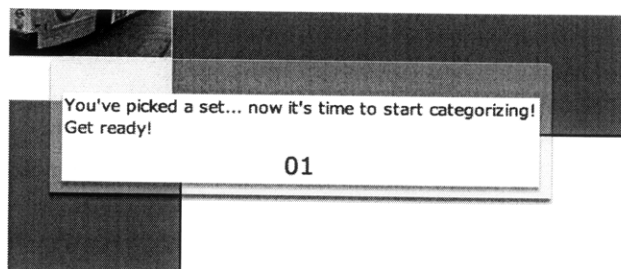
Figure 7: The countdown to the category selection screen.

The images are loaded at this point, because at this point, the host computer sends the image set information to all the other players. For smaller sets and for computers with faster connections, the image loading is near instantaneous, but for the larger sets and for slower computers, the image loading may be a long process. For this reason, we decided to add in an image loading progress bar that blocks the rest of the game while the images are loaded.
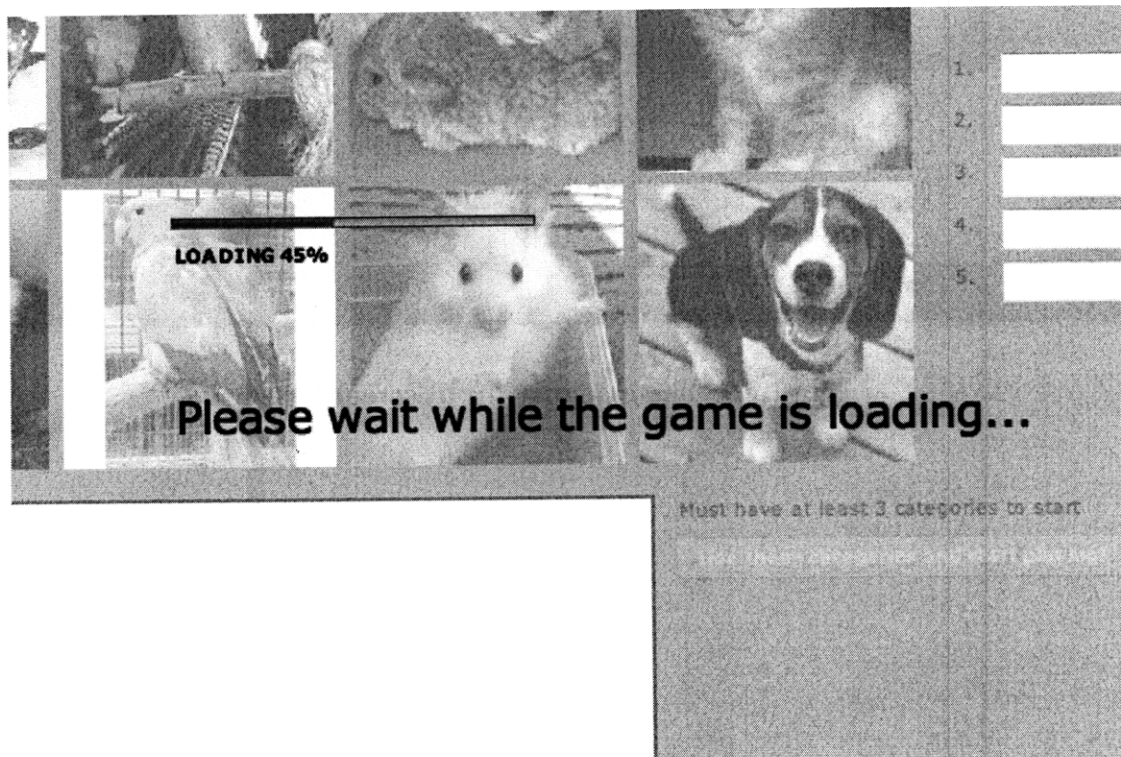

Figure 8: The image loading progress bar.

After the image loading process has finished loading, the categories can be entered in. As the image loading time differs for everyone the timer is not started

until the last player has finished loading the images. At this time, the category
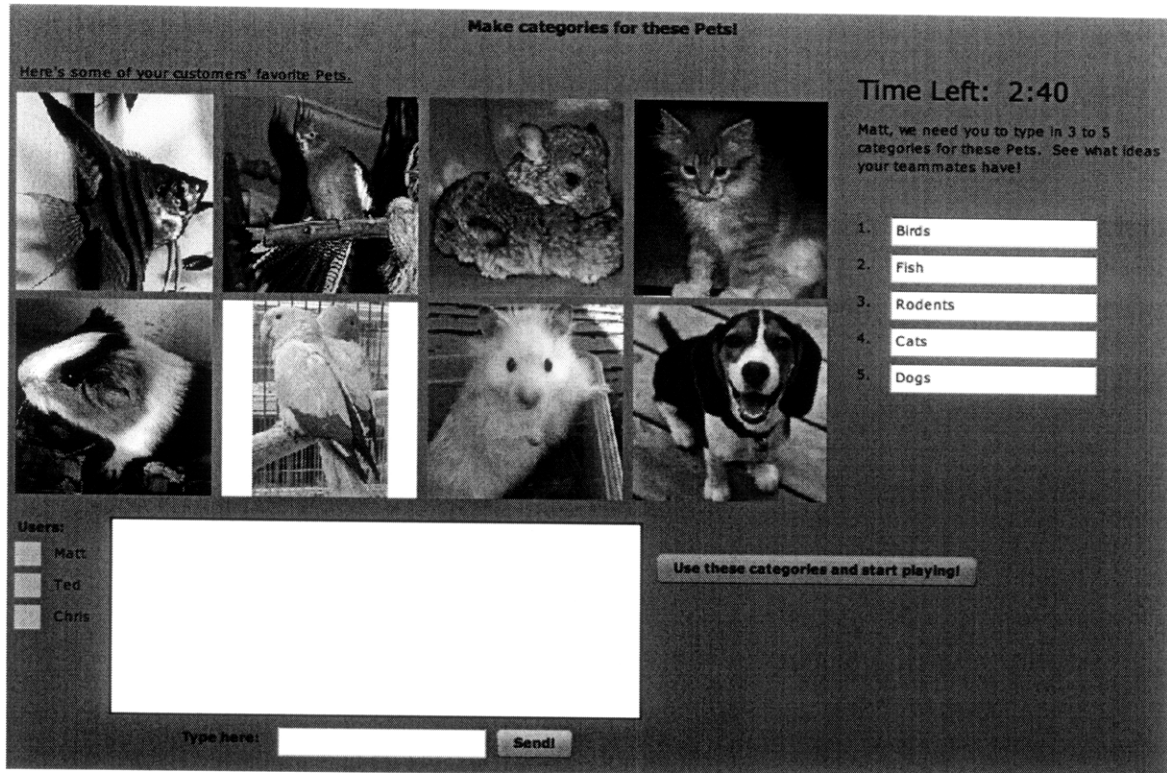
selection process can start.

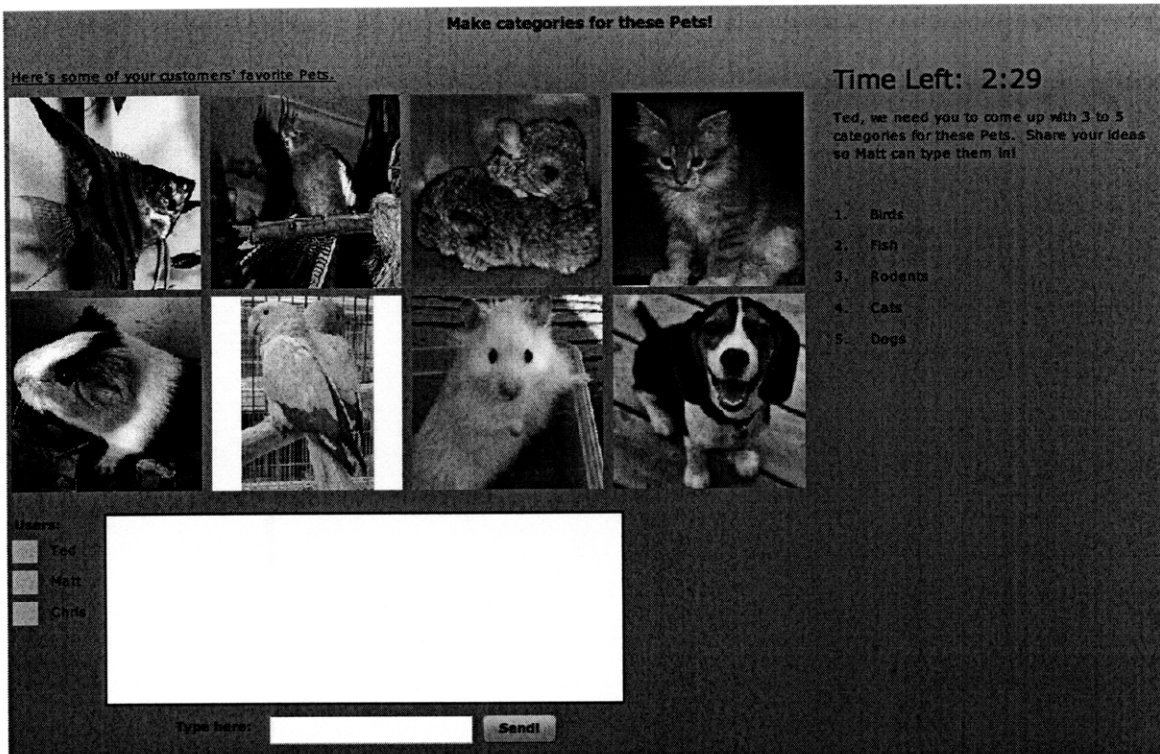**Figure 9: The category input screen.**

Figure 10: The category display screen.

Most of the players are shown the screen in Figure 10, but the host player is given the screen in Figure 9. The only difference is that the host player is given the ability to input the categories, and the other players are shown the list of categories, which is updated with each character changed or added by the host player in real time. The chat is featured as an essential part of the screen as it is used for the discussion of what the categories should be. Note how the host screen and the non-host screens both have messages that encourage dialogue between the players. For consistency and fairness, all the pictures shown are the same, so the discussion can fully focus on the pictures and what categories would be best. The total time for discussion is 3 minutes long, and this stage of the game ends when either the time runs out or when the host player clicks the "Use these categories and start playing" button.
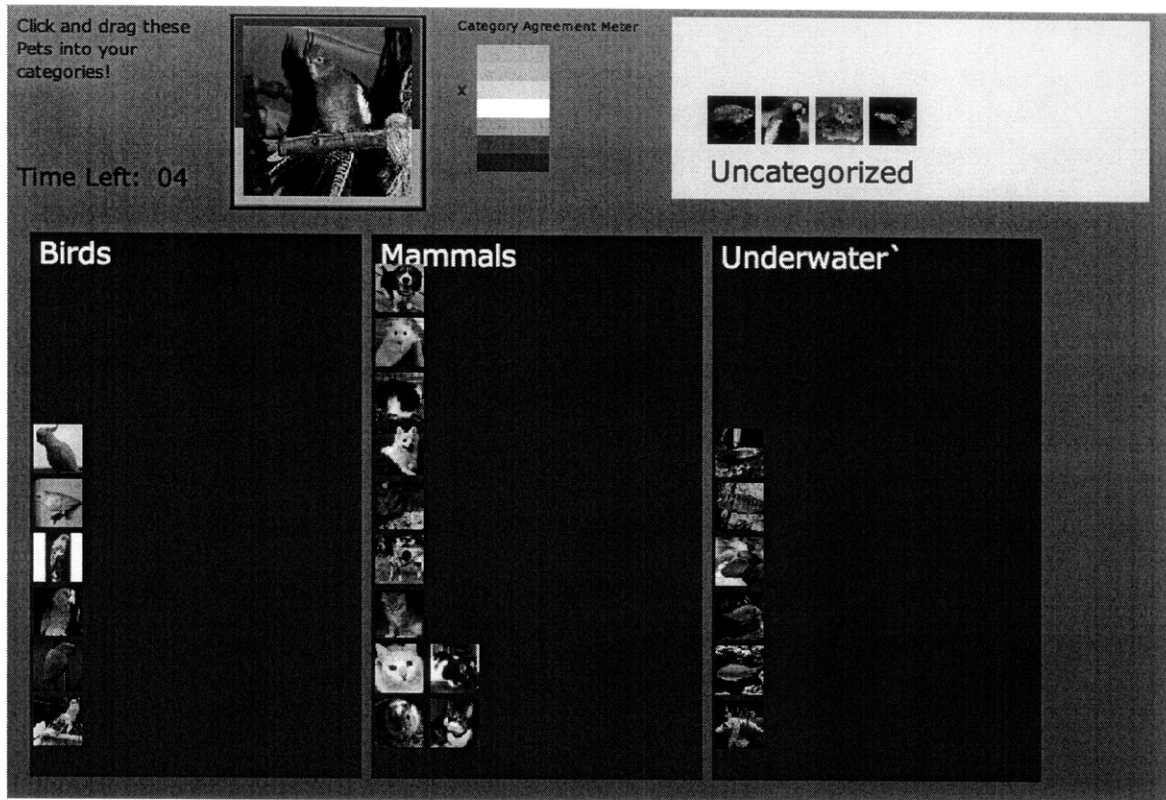
Figure 11: The drag drop categorization screen.

The players are taken to the drag drop categorization screen, where the players are shown the images one at a time, and they are also presented with a box for each category that they can drag and drop the images into. The message on the upper left informs the players that they can click and drag the images into the categories, which are represented by the large boxes. Once an image is dropped in one of the categories, the image shrinks into a smaller size and is snapped into place into an appropriate spot in the category's box. The program automatically aligns the images into columns of increasing height. A shrunken image can be momentarily enlarged by clicking on it inside the box. When in play, there are 2 visual indicators that alert the player to the time limit. Firstly,

there is a timer on the upper left side of the screen. Secondly, and more apparently, there is a dynamic border around the image to be dragged. The border starts completely green in the beginning of the period, and as time ticks on, the border becomes more and more red. When time runs out and the image has not been categorized, the border becomes completely red, and the image is placed in the "Uncategorized" box. In the case of the time running out or the image being categorized by the user, a new image is immediately presented, ready to be categorized, and the timer is reset. Also at the top of the screen is an agreement meter, ranging from green (meaning good) to red (meaning bad).
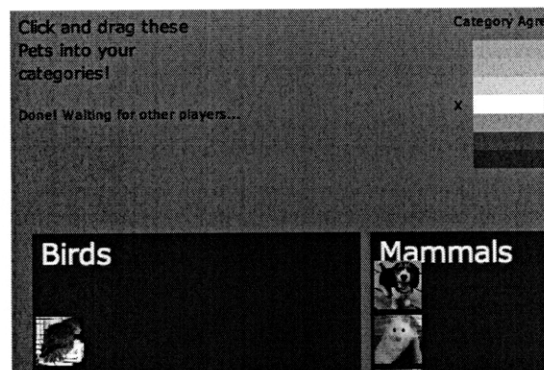

Figure 12: The message to a finished player.

After all the players have categorized a subset of 8 consecutive images, the agreement level is calculated and the meter gives feedback as to how well the players are doing. When a player has finished categorizing all the images, he or she is presented with a message saying that they are done and need to wait for the rest of the players to finish. After all players have finished, they are brought to the results screen.
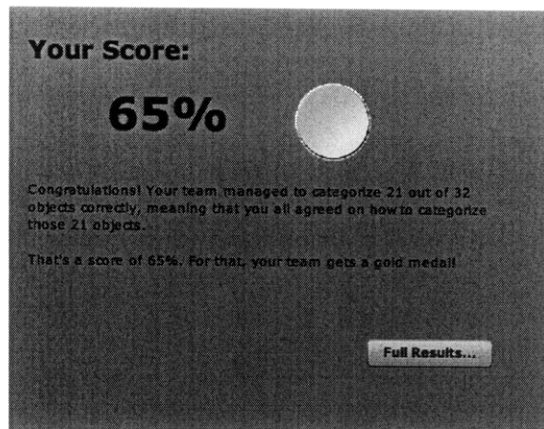
Figure 13: The initial results screen.

The initial results screen displays the percentage of images that the team categorized correctly, meaning the images where the team members were in full agreement. With different percentages, different prizes (and corresponding images) are awarded to the team with more exciting and extravagant gifts for better performing teams. The full results are shown when the player clicks the "Full Results..." button.
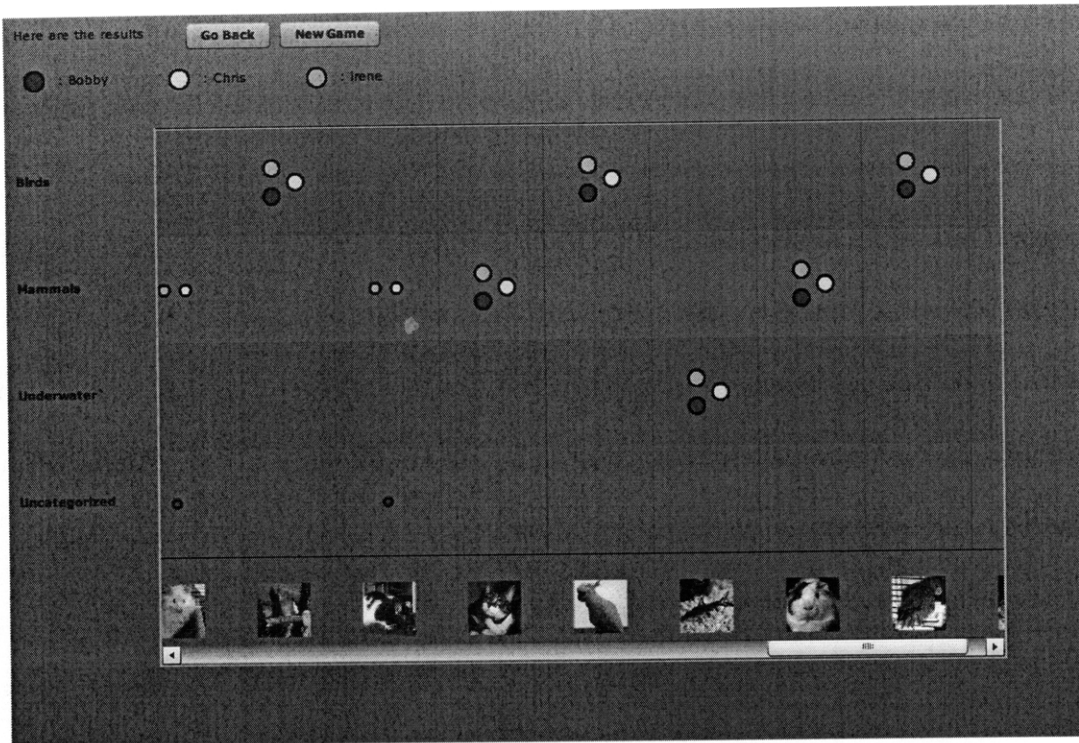
Figure 14: The full results screen.

The full results screen is organized in a grid like manner, and displays how every player categorized every single image. Each category created by the team has an entire row in the grid (including "Uncategorized"), and every image is given a column. The actual image of each categorized image is shown on the bottom of the column, giving strong consistency to the game. Every player is assigned a color, and for each image column, a dot of the player's color is displayed to show how he or she categorized that image. To stress the importance of agreement, the dots are bigger when more players agree on a categorization. The results screen was implemented to both show how each player categorized and also to emphasize how often the players agreed and how well they chose good separating categories.

5. Conjecture Graph Spoons

Conjecture Graph Spoons (or *Graph Spoons* for short), is the second multiplayer online game created for Kids' Survey Network. The design staff of The Education Arcade, led by Caitlin Feeley and Scot Osterweil, designed the game. The implementation was done by Dana Tenneson, Colin Greenhill, and myself.

5.1 Concept

The objective of *Graph Spoons* is to familiarize students with the idea of visualizing data, namely the concepts of hypotheses and conjecture graphing. The game is a multiplayer computer-based card game where the players attempt to assemble a "winning hand" of cards that contain the various parts of a conjecture graph for a given hypothesis. In addition to hypotheses and conjecture graphing, the game also stresses the understanding of graph terminology, graph interpretation, as well as peer review and feedback.

The deck used in the game play consists of 3 different kinds of cards; graph shape cards, label (variable) cards, and value (range/category) cards. In each game session, there are 6 hypotheses given to the entire set of players, and the deck has suitable cards to create correct conjecture graphs for the hypotheses. Each player has a dynamic workspace where he or she can drag the cards to create a representation of the conjecture graph that he or she is trying to create. There are 4 different kinds of graph shape cards: line graphs, bar graphs, scatter

plots, and pie charts. The shape cards show some graphical representation of a prediction without any of the labels or values that would specify the details of the graph; as the name implies, they are merely the shape of the graph. Clearly, only some of the label and value cards would make sense for certain graphs. The label cards identify the measurement on either the independent or dependent axes with titles such as "Number of Science Club Members." The value cards designate the range of the data displayed on the graph, thus adding numbers to the axes of the graph. The value cards usually have numerical values, but sometimes they have nominal ones to designate categories for pie and bar charts. In addition to the standard cards, there are also wild cards and effect cards. Wild cards can be used as any of the other standard cards on the deck, selectable by the player. Effect cards are special cards that allow the player to engage in special game actions such as steal a card from another player.

The game play of *Graph Spoons* is loosely based on its namesake, the card game of Spoons. Each player starts with 5 cards in their hand, which can be of any type. The game is made up of multiple rounds, and the "dealer" changes after each round. In the beginning of each round, the dealer draws a card from the deck, giving him or her 6 cards. He or she then picks one of the 6 cards to discard to the player on his or her right. The next player then chooses a card to discard to the player to the right, and the round continues until every player has received and discarded one card. A timer is put in place to ensure that each player discards a card in a timely fashion. If the timer runs out, a random card is discarded for them. This keeps the game going at an enjoyable pace.

At all times, the players can create a conjecture graph by moving the cards onto spots designated for value label or shape cards. Placing cards on these

designated drop zones adds the corresponding value, label or shape to the graph displayed for that user. Once a player has created a graph that they think matches one of the hypotheses, they can submit the graph. At this time, the submitted graph and hypothesis selections are sent and displayed to all the other players. The players either approve or disapprove of the submission, and are awarded or deducted points for being correct. The submitting player either receives a much higher amount of points for submitting a correct graph or loses a higher amount of points for submitting an incorrect graph.

The game strongly reinforces many of the learning goals. To succeed in the game, the players must learn how to create a conjecture graph: a difficult task involving understanding hypotheses, graphs, and variables. The players receive hands-on experience in creating conjecture graphs by physically moving cards into spots to create a visual representation of a hypothesis. The different cards split up the various parts of the graph, which emphasizes the importance of each part. In addition, the review of the graph submission by all the players forces them to critically analyze a graph and decide whether it matches a hypothesis or not.
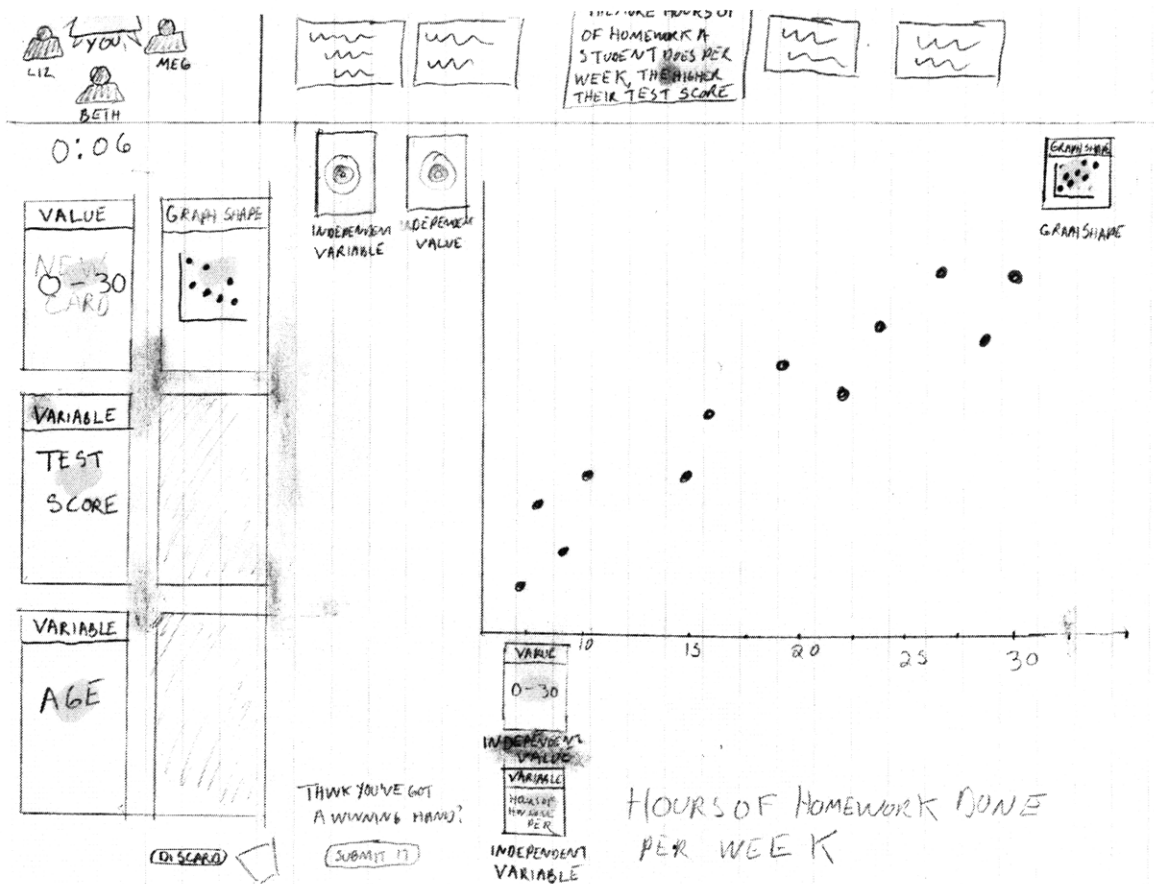
5.2 Front End Design

Figure 15: The paper prototype for Conjecture Graph Spoons.

The design of *Graph Spoons* was first sketched out on the paper prototype shown in Figure 11. In the paper prototype, the main focus is given to the actual graph assembled by the player according to the large amount of space allocated to the graph. Alongside the graph are the spots to place cards that apply the values onto the graph. The top left of the screen features a small box where all the players are represented with an icon. That box conveys information such as whose turn it is, who the dealer of the round is, and how the cards are being distributed. To the right of that box is the hypotheses area. The hypotheses are

lined up in hypothesis boxes, where the selected hypothesis is enlarged and highlighted. The playing cards that are not put in use in the graph are stored in the card hand area to the left. With our implementation of the game, we stayed fairly close to this early design, making a few changes for usability on the computer screen.

5.3 Back End Design

The format of the *Graph Spoons* game led to a fairly direct object oriented data representation for the game. The abstract class, Card, is the basis of all the various kinds of cards that can be played. Besides storing basic information such as id numbers, the Card class' most common usage is to facilitate polymorphism with the various specific card classes and to also provide a equals() method that allows the comparison of two cards. The Card class is extended by its subclasses ValueCard, VariableCard, ShapeCard, WildCard and EffectCard. ValueCard, VariableCard, and ShapeCard store the appropriate text or image and graph data that represents what the card is. WildCard can represent any other card, so it stores a reference to a card that it could represent. All the cards are stored in a Deck object, which has the ability to shuffle the cards and to deal a card. Additionally, Hypothesis objects are stored inside the Deck, as a set of cards have a matching set of hypotheses they build around.

In addition to the room creation and joining server code similar to *Categorize This, Graph Spoons* takes advantage of the Smart Fox Server's messaging and room variables to handle a lot of the communication and consistency between the different players of the game. Real time updates of the

game state, events, effects, scoring, and shuffled decks are sent between the players to ensure uniformity between the players. Communication is essential during the portion of the game when a player submits a graph. The entire graph and chosen hypothesis are sent to every other player, so that each player can verify the correctness of the graph. Lastly the total points of each player are sent to everyone else at the end of the game.

5.4 Implementation

Similar to *Categorize This*, *Conjecture Graph Spoons* is currently in a beta mode where the playability is in place but final artwork and visual design are not included at all. The game was designed fairly close to the paper prototype design, with changes made to best utilize the limited pixels available on a computer screen. The game play—including the basic graph creation, card passing, and events such as effect cards being played—are all implemented and ready to use.
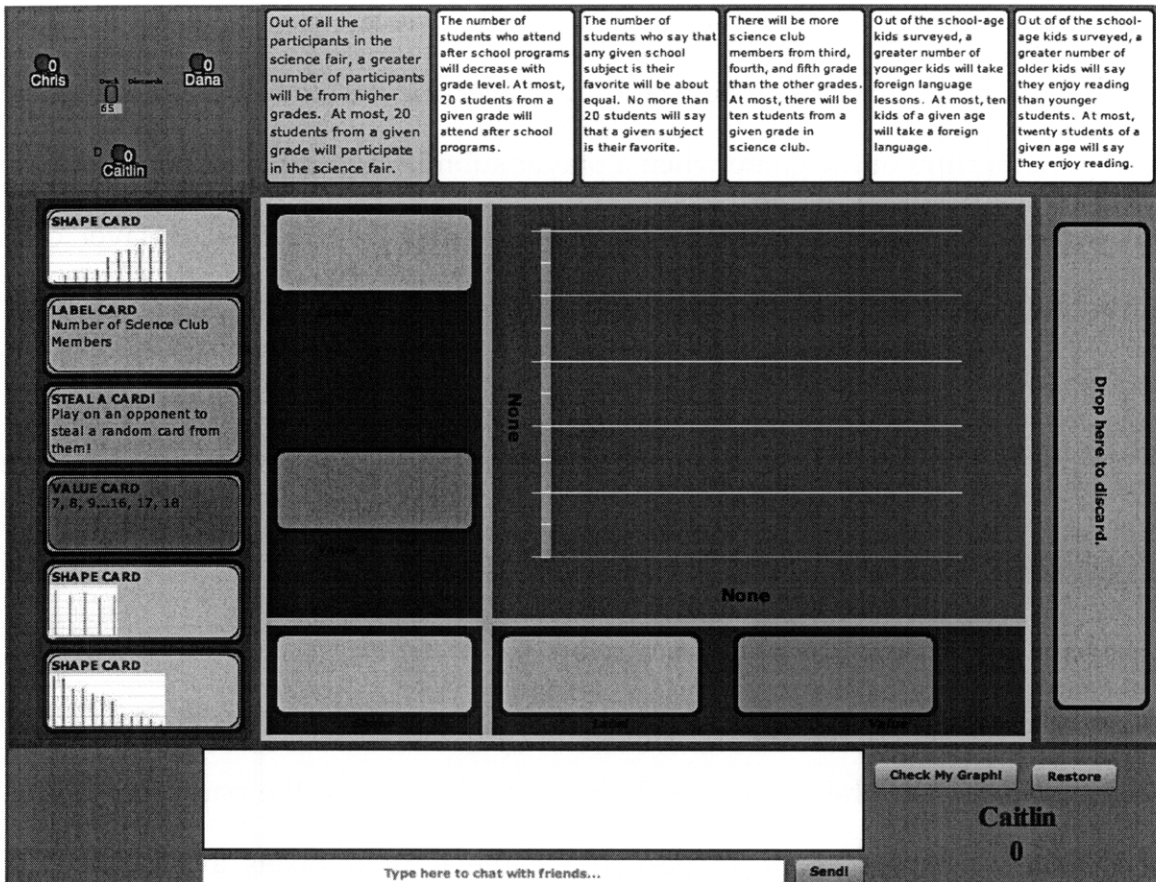
Figure 16: The full screen of Conjecture Graph Spoons.

Figure 16 shows the screen of *Graph Spoons* at the start of a game. The screen can be broken up into a few main parts. The top of the screen is split between the player's box and the hypothesis area. The middle section, largest in area, contains the player's hand area, the graph area (the largest and most featured section), and the discard area. The bottom control panel contains a chat box and buttons to submit the graph and to change the view settings.
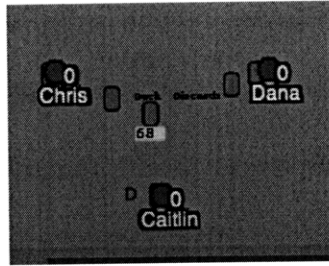
Figure 17: The player's box's card dealing animation.

The player's box, in the top left of the screen, displays all the players in the order they are "seated" around the table and also the number of cards in both the deck and the discard pile. The player whose turn it is has his or her name highlighted, and a "D" designates who the dealer is. As cards are dealt from the deck or passed from player to player (or into the discard pile), little card animations show the card movement inside the player's box.
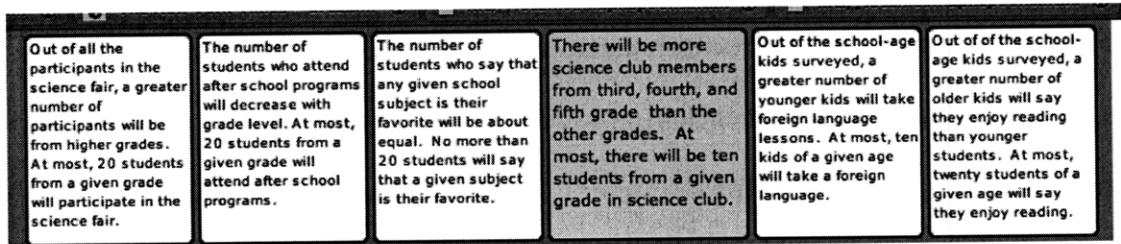

Figure 18: The hypothesis selection area.

Despite the similarity of the finalized version to the paper prototype, the hypothesis selection area actually changed significantly in appearance during the many iterations of development. Earlier versions had the selected hypothesis taking much more screen space, with a larger font for a more readable selected hypothesis. However with that version, it was impossible to fit the complete text of the other hypotheses on the screen in a practical fashion. Thus, a key decision

was made to make all text smaller but visible on the screen at all times. The selected hypothesis box (selected by clicking) is highlighted in bright yellow, and the box size (and font) is slightly larger than the rest, which clearly differentiates the selected hypothesis against the others. Resize animations are done with each click to increase the user satisfaction of the game, and to increase the memorability of the hypothesis selection process.
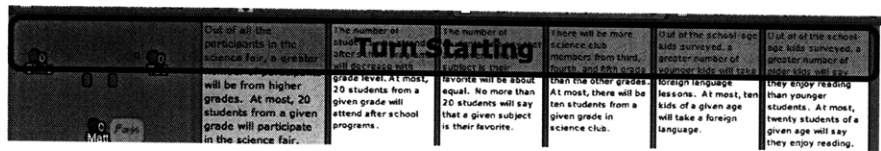

Figure 19: A "Turn Starting" message presented to the player.

To emphasize events occurring in the game, messages drop down to tell the players of things that are happening. These messages are extremely important, as many events would happen without the player knowing otherwise. These events could be an effect being played, which could lead to a player's card being stolen or having their turn skipped. The "Turn Starting" message is also important, as it ensures that the player realizes that it's his or her turn, which reduces long delays of wasted time in the overall game play.
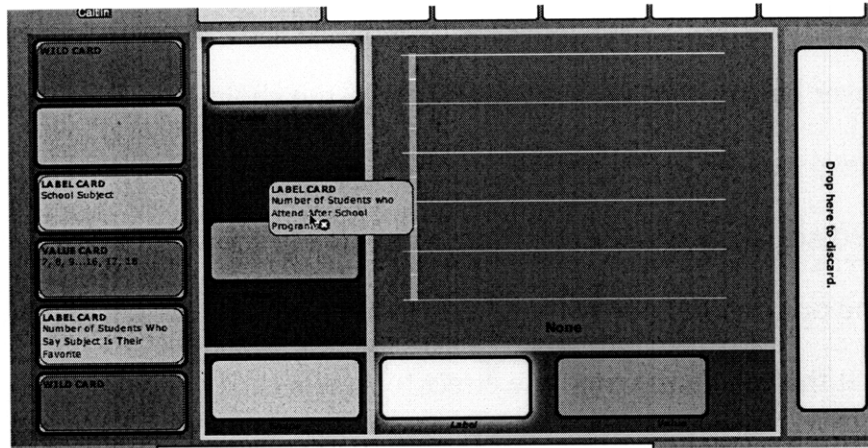
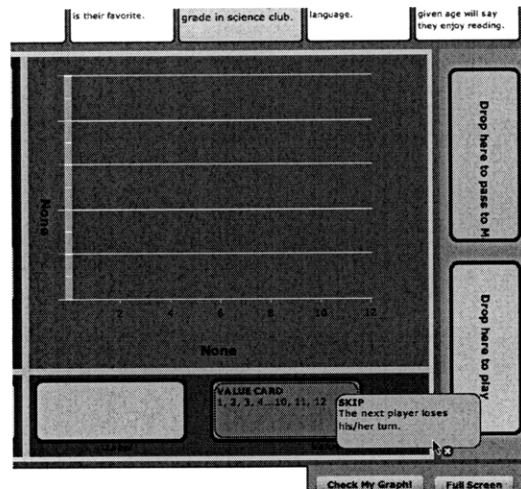**Figure 20: The click and drag of a Label Card.**

Figure 21: The click and drag of an Effect Card.

Figure 20 shows the click and drag capabilities of the game. From the hand area on the left, cards can be clicked and dragged into other areas on the screen. Visual cues highlight the valid areas where the cards can be dropped. The suitable drop zones flash brightly and also have a glow effect around them. In addition, a red "stop" circle is presented at any times where releasing the drag would result in returning the card back to the hand. When a Wild Card is dragged, all of the card zones flash and glow, as they are all valid drop zones.

When a Wild Card is placed on a drop zone, a panel pops up to present the choices the user has for that spot. Appropriately, the choices are all the possible values for the given card spot. Lastly, when an Effect Card is clicked and dragged, the discard/pass area is split into two areas, shown in Figure 21, as the player has the option of either playing the effect or passing the card. Figure 22 shows a panel that pops up when the effect to steal a card from another player is played. The panel presents the user with the possible cards that he or she can steal from another player.
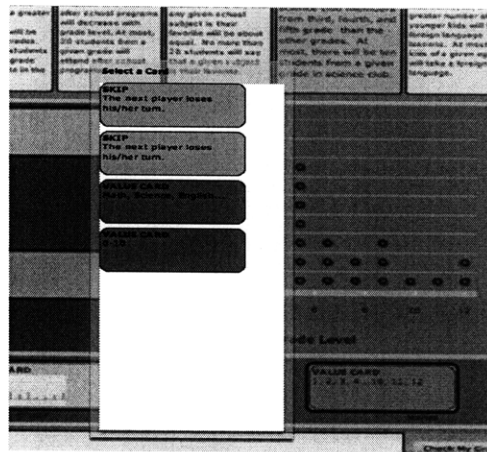

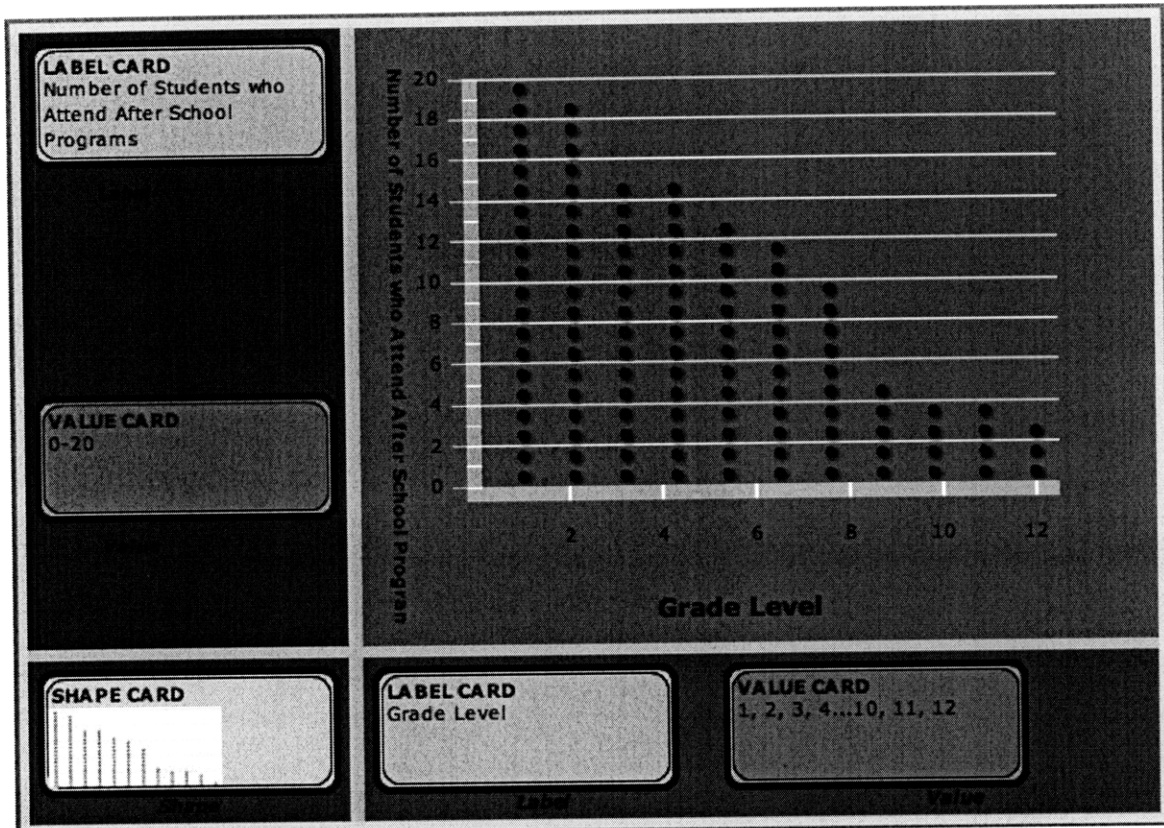**Figure 22: The card stealing effect pop-up.**

Figure 23: An assembled graph in Conjecture Graph Spoons.

As cards are dragged into their appropriate drop zones, the graph area changes to accommodate the added information, such as the labels or the graph shape. Figure 22 shows a fully assembled graph showing relationships between the number of students who attend an after school program and the students' grade level. Appropriate Value Cards were needed to give values on the axes that match the Label Cards. The Shape Card features a histogram that conveys the idea that there are many more students in lower grade levels who are also in after school programs.
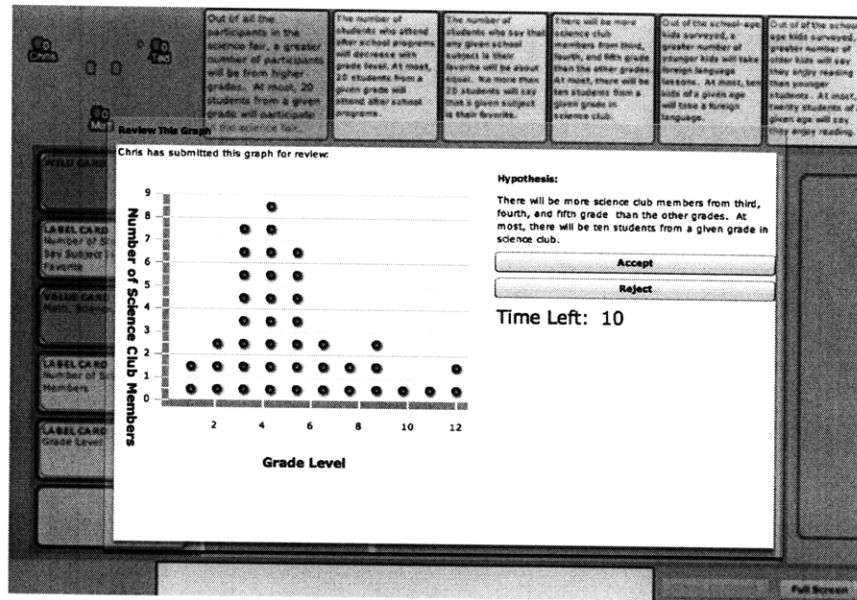
Figure 24: The submission approval screen.

When a player thinks that he or she has successfully assembled a graph that matches with one of the hypotheses, he or she submits the graph. After a quick dialog to review the graph and hypothesis, the conjecture graph is sent to all of the other players. The submitting player receives a large amount of points for submitting a correct graph, but the other players receive points for correctly accepting or rejecting another player's submission. This matches with the learning goals of being able to read and understand a conjecture graph and building a community of peer review. The game may be finished after the first player successfully builds a conjecture graph, or it may continue for multiple rounds, depending on the options the players have chosen. When the game ends, the players are presented with the final screen that gives an overview of the last submitted graph and shows the final points score totals among the players, shown in Figure 25.
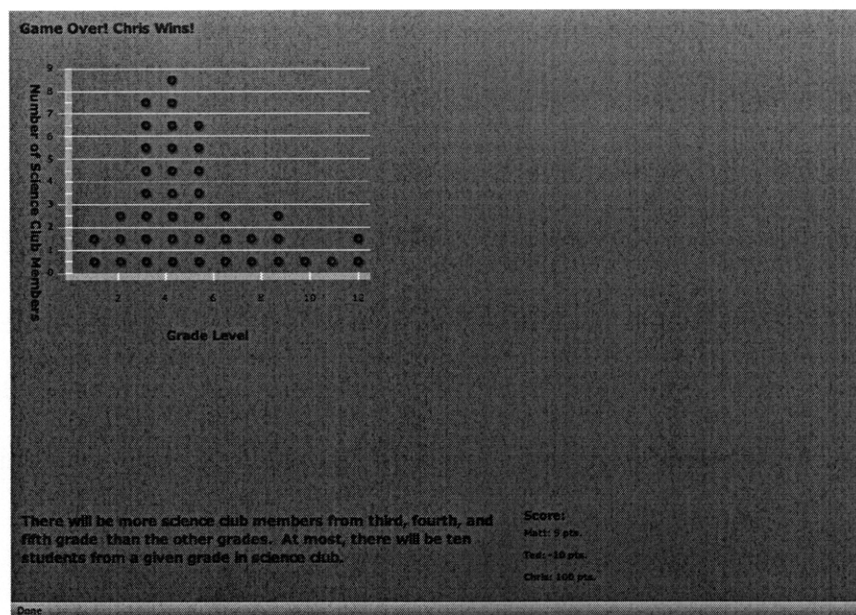
Figure 25: The final screen of the game.

Overall, the game takes advantage of the instant computation, networked communication, and animation to create a fast paced, exciting game that strengthens the players' understandings of conjecture graphs. The students must consider multiple hypotheses while contemplating numerous graph shapes, labels and values. The additional effect cards and wild cards add both excitement and critical thinking to the game. For these reasons, *Conjecture Graph Spoons* is an effective game for teaching conjecture graphs.

6. Testing Results

Pilot testing was done for both *Categorize This* and *Graph Spoons*. The testing brought to our attention both good and bad aspects of the games. Some of the input led to immediate changes in the games, and the experiences also provided information on how the players achieved the learning goals during the game.

6.1 Categorize This Pilot Testing

The *Categorize This* pilot testing occurred in December 2008 with testing subjects from Millis Middle School in Millis, MA. There were three hour-long testing sessions where the students would play the game and verbally comment on the experience while playing. After multiple rounds, the players were asked various questions to prompt responses and discussion.

The players were generally very enthusiastic about the game and had mostly positive reactions to the game. The students were especially enthusiastic when the image sets were related to topics that they liked (the girls in the test called the shoe images their favorite) and also when it was their turn to be team leader, as they acted with increased authority when they were the leader. The students recommended more category sets such as toys, famous people, animals, and sports. They also recommended more stimulating imagery and dynamic animations as the lack of audiovisual feedback from the game made the game slightly boring and unsatisfying. A very beneficial response with respect to the

learning goals was that most of the players found the concept of categorization and sorting very enjoyable.

In general, the set selection, category selection, and sorting sections of the game created excitement and discussions about the categorization without any prompting or input from the administrators. The players' abilities to select well-defined and appropriate categories improved over time. At first, the players created excessively specific categories which led to the later realization that many items did not fit into any of the categories they had written. From that experience, the players learned to create meaningful, mutually exclusive categories –a learning goal of the game. The players also realized that their categories often had problems with specificity. For example, with the food items image set, they created categories for lunch and dinner to only realize that many food items were eaten in both lunch and dinner. The improvement the players made in terms of picking categories showed that they also progressed in understanding inter-rater reliability, outliers, and prediction. The one aspect of the game that needed administrator assistance was to encourage the players to deeply examine the results screen. The players seemed to quickly skip over the results screen as they wished to play the next round. It was not until the supervisors encouraged them that they began to examine the results to see where they went wrong.

As a result of these experimental testing sessions, we compiled a list of revisions that we considered while continuing the implementation of the game. Many of the revisions were of the actual content of the game, including ideas for image sets and difficulty settings. The size of the images was increased to allow

for better viewing of the images. The agreement meter was made to update more frequently, for better granularity of the game. The placement and size of the timer and category labels were changed to make both items more prominent. An "I'm Ready" button was replaced by a countdown timer in between the different stages of the game, and the results screen was slightly reworked to emphasize the meaning of the colors. These changes were implemented into the current version of the game, while ideas such as sounds and animations were designated to be completed with the final implementation and artistic interpretation.

6.2 Graph Spoons Pilot Testing

There has not yet been testing of the implemented version of *Conjecture Graph Spoons*; the only testing has been with paper prototypes consisting of real paper cards in a real deck. A paper game is not as visually engaging nor is it as fast paced as a software version could potentially be. However, we received a significant amount of ideas and feedback about the game from the testing. Predictably, the main input from the players was related to making the game more fun to play. The basic graph-building portion of the game was deemed too school-like and unexciting, and a group of additions were thought up, including the Wild Card and the Effect Cards. The unpredictability and competition were emphasized in the new additions to the game.

Good signs of student learning were also prevalent in the observations of the testing sessions. In addition to the players being successful in creating graphs to win, the players were also observed making intelligent decisions during the game, such as choosing not to discard cards that they thought the other players

wanted. Overall, the game seemed to have a significant amount of potential to be

a fun and effective game to teach conjecture graphs.

7. Future Work

The future work that we need to do with the Kids' Survey Network includes the creation of additional games, the finalization of the artwork, animation, and sounds with the games, and also the integration of the games with the larger Kids' Survey Network site created by TERC.

The third game, *Superhero Draft*, is currently in the middle of production. The premise of the game is that each player is an employee of a different superhero agency, and they are each given the task of building a team of superheroes from a common pool of squads. For each particular round, there is a different dangerous situation threatening the city, which may range from a natural disaster to a monster attack. For each particular case, there is a different ideal distribution of superhero power levels for a team of superheroes. The threat introduction screen is shown with some sample text is given in Figure 26.
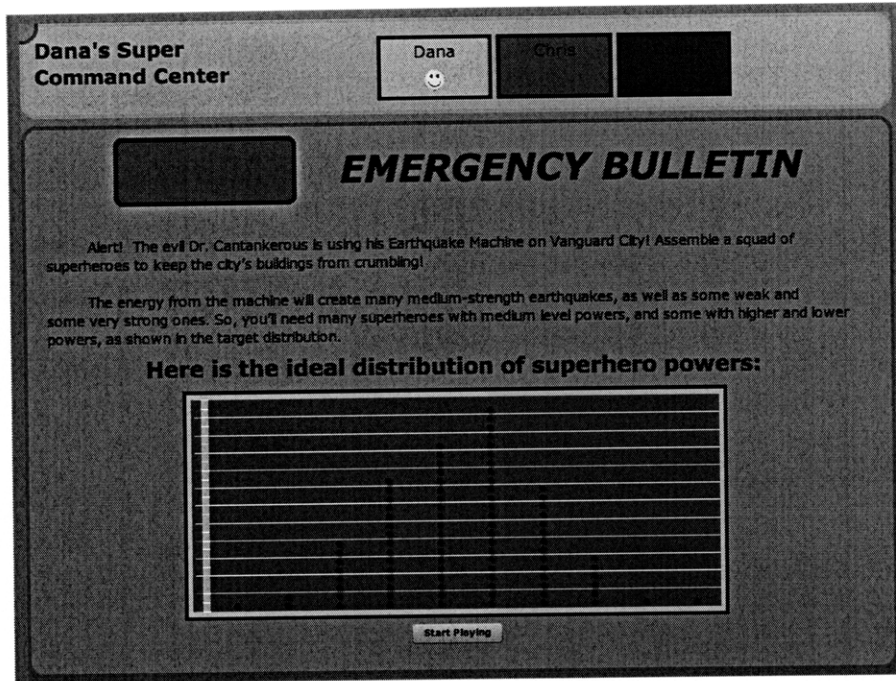
Figure 26: The Threat Introduction screen of Superhero Draft.

During the draft, each player chooses three squads, each with their own different distribution graphs. After the draft, there is a trade period where the players can choose to trade squads amongst each other. The player's final distribution is created by directly adding all 3 of their finalized squads to create one complete distribution. This final distribution is compared against the target distribution to compute the outcome and final score, stressing important topics such as mean, median and range. The results screen is shown in Figure 27.
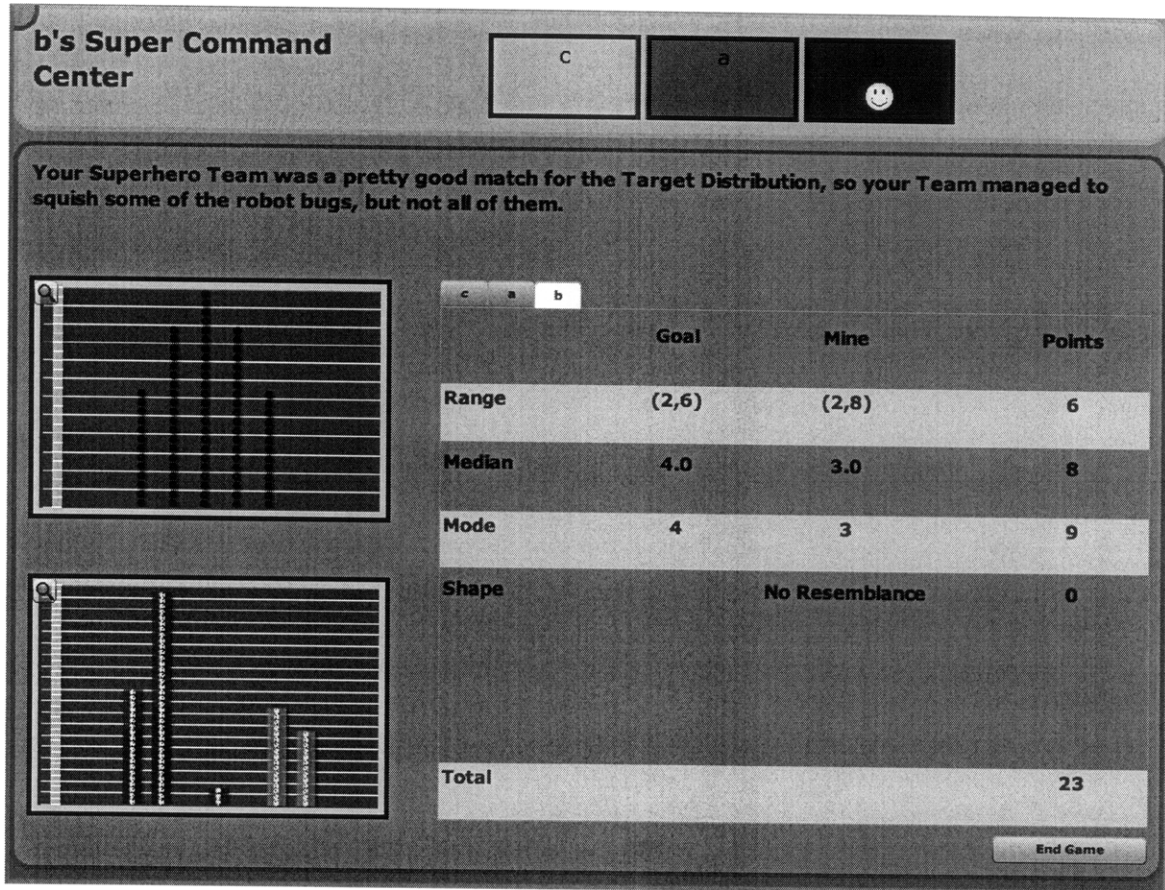
Figure 27: An early stage results screen for Superhero Draft.

Superhero Draft forces its players to think critically about graph distributions, especially how differently skewed distributions that include outliers add up to a completely different larger total distribution. The concepts of range, median and mode are also introduced and reinforced during the score display stage. Since the players are forced to examine many different smaller distributions during the draft, they must reflect on these concepts many times during the game. The implementation of *Superhero Draft* is nearing completion, and the entire project is ahead of the timetable.

8. Conclusion

The stated goal of Kids' Survey Network is to generate an apprenticeship network, a project based community where students learn concepts of data literacy. We have succeeded greatly in our task of designing and implementing the suite of multiplayer games that compliment the larger Kids' Survey Network. We have created games that successfully teach concepts about surveys, data, and graphs. The games are entertaining and exciting to play, thus producing an attention-holding educational platform. Lastly, the multiplayer aspect further builds interactions and a community mood among the participants, as the games all include multiplayer-specific game play. With the finalization of the rest of the Kids' Survey Network project, we are confident that our work will make a significant positive impact on the educational levels of the project's subjects.

Bibliography

[1] Andee Rubin, "Math That Matters," *Threshold Magazine*, pp. 22-31, Spring 2005.

[2] Paul Graham, "Web 2.0," November 2005. [Online]. Available http://www.paulgraham.com/web20.html

[3] Daniel Terdiman, "Video game industry posts big February gains," March 2009. [Online]. Available http://news.cnet.com/8301-10797_3-10200720-235.html

[4] Kurt Squire, "Cultural Framing of Computer/Video Games," July 2002. [Online]. Available http://www.gamestudies.org/0102/squire/

[5] Tiffany Wang, "Case for Usability: Designing Outdoor Augmented Reality Games," August 2008. Master's Thesis, Massachusetts Institute of Technology.

[6] "Terc: About Us," 2006. [Online]. Available http://www.terc.edu

[7] "Kids' Survey Network: Developing and Studying an Apprenticeship Network for Informal Math and Science," 2009. Project Proposal, TERC, Inc.

[8] "Adobe Flex 3," 2009. [Online]. Available http://www.adobe.com/products/flex/

[9] Tariq Ahmed, "What is Adobe Flex," 2007. [Online]. Available http://www.cflex.net/about_adobe_flex.cfm