Children's Learning from Multiple Media in Informal Mathematics Education

Shalom M. Fisch
MediaKidz Research & Consulting

Richard Lesh, Elizabeth Motoki
Indiana University

Sandra Crespo, Vincent Melfi
Michigan State University

Advisory Board

Research Team

Janet Alcini
Meredith Bissu
Kristy Brugar
Kim Cervello
Belinda Chiu
Dan Clark
James Colee
Candice Collins
Ed Corcoran
Cindy Covell
Corvell Cranfield
Sandra Crespo
Hanna Figueras
Shalom Fisch
Anny Fritzen
Funda Gonulates
Huang Hsuan-Yi
Kate Johnson
Jerilynn Lepak
Richard Lesh
Lorraine Males
Carmina Marcial
Susan McCann
Jeanne Meier
Vincent Melfi
Elizabeth Motoki
Joy Oslund
Sam Otten
Maria Selena Protacio
Jennifer Shulman
Nava Silton
Faith Smith
Marj Terpstra
Justin Thorpe
Carolyn Volpe
Rachel Wiesner
Linda Young

Acknowledgements

This research was funded by a grant from the National Science Foundation (DRL-0723829). We gratefully acknowledge the entire *Cyberchase* production team at Thirteen/WNET (especially Sandra Sheppard, Frances Nankin, Michael Templeton, Rekha Menon, and Cathy Cevoli) for their support, and online producers David Hirmes and Brian Lee for building the tracking software used here. We also thank our program officers at NSF, Sue Allen and Elizabeth Vanderputten, and – most of all – the staff, teachers, and students of the participating schools for their cooperation and patience throughout a long and intensive study. Without them, the research would have been impossible.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Many informal science and mathematics education projects employ multiple media, but studies typically have investigated learning from a single medium, rather than multiple media. The present research, funded by the National Science Foundation, used *Cyberchase* (a multiple-media, informal mathematics project targeting 8- to 11-year-olds, produced by Thirteen/WNET) to investigate synergy among multiple media components and how they interact to yield cumulative educational outcomes.

A total of 672 children, in nine public elementary schools in Michigan and Indiana, participated in the study. The research incorporated both naturalistic and experimental methods, to investigate children's use of *Cyberchase* media, ways in which use of one medium feeds into use of another, and the educational impact of *Cyberchase* on children's problem solving and attitudes toward mathematics.

The study was designed to address the following research questions:

1) How does the mathematics learned from multiple media differ from mathematics learned from a single medium?
2) What outcomes derive from engagement with different types of media, and what types of synergy occur?
3) How can reliable research methods be developed to assess contributions of individual media and their interactions?
4) How can informal education projects capitalize on the strengths of each medium?
5) How can media components be designed and employed to best complement each other?

Highlights of the results include the following:

Patterns of Naturalistic Use

- Use of each form of *Cyberchase* media (TV and Web site) was fairly consistent over time. Those children who watched *Cyberchase* on TV in one month tended to do so in subsequent months as well. A similar pattern was found for month-to-month use of the Web site.

- Children's use of *Cyberchase* also tended to span media; each month, children who watched the *Cyberchase* TV series more frequently also tended to visit its Web site more often. Thus, in naturalistic use, some children do indeed use multiple media when they are available (which lends real-world validity to the question of how children learn from multiple media).

- Because most users' first encounter with *Cyberchase* occurred long before we began collecting data, the present data cannot determine which medium came first. However, past research found that children more often begin by watching the TV series and subsequently expand to using the Web site as well.

Learning from Cyberchase

- Past research (which evaluated the educational effects of the *Cyberchase* TV series alone) found evidence of significant impact on both the process of children's mathematical problem solving and the sophistication of their solutions. The present study replicated that finding, and extended it by finding more consistent effects of video plus online games than of either medium alone (especially in comparison to online games alone). Interestingly, learning from *Cyberchase* was not manifest in children's simply doing a greater number of things while working on the tasks, but rather in their using a greater *variety* of strategies and heuristics, and in using those strategies and heuristics more effectively. In addition to quantitative, statistical comparisons, qualitative observations revealed that children demonstrated persistence and top-down planning while working on tasks in the posttest. As in past research, effects emerged more consistently in tasks about organizing data (e.g., combinatorics, predicting from data) than in tasks about measurement.

- Surprisingly, however, children in the DVD + Web group also showed consistently greater gains than children in the All Materials group (which used the same materials plus hands-on classroom activities). Although we cannot be certain, we believe that the less consistent performance of the All Materials group may have been influenced by cues from teachers in response to the demands of having to make time for *Cyberchase* media or hands-on activities every day.

- Effects on problem solving often appeared to be driven more by the TV series than by the online games. We suspect that this is due to the fact that television is designed to serve as the central component of *Cyberchase*, and provides greater explanation of mathematical concepts than the games (which allow children opportunities to exercise skills, but present less overt explanation). Such explanation -- embedded in the context of appealing characters and a compelling narrative -- appeared to provide both the necessary understanding and modeling of processes and dispositions for effective problem solving (e.g., persistence, top-down planning). However, games designed for more overt instruction and explanation (e.g., via online agent characters who scaffold children's performance) might produce stronger effects of their own.

- Although the television series produced stronger pretest-posttest effects than the online games did, online tracking data indicated that the games provided a context for children to engage in rich mathematical reasoning – and that this process of reasoning was detectable, not only through in-person observations, but also through data mining of online tracking data. Parallel to prior research on formal classroom mathematics, children engaged in cycles of increasingly sophisticated mathematical thinking over the course of playing an online game, with shifts in strategies indicated by predictable patterns of responses, such as clusters of errors or use of a "clear" button to try again.

Multiple-Media Learning

- Data on children's performance while playing the online games revealed evidence of transfer of learning, not only from the treatment to our posttest measures, but also from children's experience with one *Cyberchase* medium to another. This points to a significant strength of learning from multiple media: The lessons learned from one medium can be applied to enrich children's experience while learning from a second medium as well.

Attitude

- Paper-and-pencil measures of attitude revealed only one pair of significant effects: From pretest to posttest, all of the *Cyberchase* groups sustained their interest and (to a lesser degree) confidence in doing school math, while the attitudes of the control group declined. No significant effects appeared for other domains of out-of-school mathematics.

- However, we also found behavioral evidence of an effect on children's motivation: In two of the three *Cyberchase* online games, users of multiple media were more likely to continue playing beyond the end of the game than children in the Web Only group, pointing to their greater motivation to engage in a fun, mathematical activity.

Conclusions and Implications

- Together, these data suggest that children use multiple related media during naturalistic use, and that such use can promote both learning and motivation toward engaging in additional, related activities for informal education. Cross-platform learning can elicit transfer of learning, both from one medium to another (resulting in richer engagement with the material) and from educational media to subsequent assessments.

- Indeed, the presence of a consistent world and cast of characters across media has the potential to serve as a bridge that not only elicits, but also facilitates, transfer of learning. In the case of *Cyberchase*, compelling narrative is used to carry both explanations of content and examples of characters who model successful approaches to problem solving, whereas participatory (interactive and hands-on) media provide opportunities for children to exercise these skills themselves. The use of a common world and characters can encourage children to connect related mathematics content across these media. At the same time, appealing experiences in one medium can stimulate children's motivation to engage in other educational activities with the same familiar characters. Over time, such experiences have the potential to stimulate interest in the embedded mathematics as well.

INTRODUCTION

Purpose of the Research

Most current informal science and mathematics education projects make use of multiple media to reach their audience. For example, one standard model is comprised of an educational television series accompanied by a Web site and local outreach. Indeed, as media platforms and delivery systems proliferate, this trend shows no signs of slowing down. Not only the production community, but public and private funding agencies, have adopted *transmedia* – the use of multiple media to tell interconnected stories – as the latest buzzword in creating children's media (e.g., U.S. Department of Education, 2010). From an educational standpoint, producers and funders assume this combination of media yields benefits for children's learning and attitudes toward mathematics or science, beyond those that might be provided by one medium alone.

However, almost no research has assessed the synergy among multiple components or explored ways in which their contributions might interact to yield a greater whole. Research on children's learning from such projects has tended to focus almost entirely on the impact of one predominant component, such as a television series or museum exhibit in isolation. In fact, a sizable research literature indicates that sustained use of educational material within a single medium can -- and does – result in significant improvement in children's understanding and attitudes (e.g., see reviews by Crane et al., 1994; Falk et al., 2001; Fisch, 2004). Yet, we are unaware of any published studies that attempt to provide comprehensive answers to the question of children's learning from multiple media, although some more narrowly defined studies have approached aspects of the issues (and some proprietary studies are not available in the research literature).

The present research addresses this gap by investigating what we shall refer to as *cross-platform learning* – that is, the potential synergy among related media components and how they might interact to yield cumulative educational outcomes.

The study is designed to address the following research questions:

1) How does the mathematics learned from multiple media differ from mathematics learned from a single medium?
2) What outcomes derive from engagement with different types of media, and what types of synergy occur?
3) How can reliable research methods be developed to assess contributions of individual media and their interactions?
4) How can informal education projects capitalize on the strengths of each medium?
5) How can media components be designed and employed to best complement each other?

Theoretical Background

Our approach to cross-platform learning is grounded in the theoretical and empirical literature on *transfer of learning* – students' ability to apply concepts or skills acquired in one context to a new problem or context. The literature has documented many different types of transfer (Haskell [2001] distinguished among as many as 14 types), and numerous theoretical mechanisms have been offered to explain how and why they occur (e.g., Gentner, 1983; Greeno, Moore, & Smith, 1993; Holyoak, 1985; Salomon & Perkins, 1989; Schwartz, Bransford, & Sears, 2005). Fisch (2004) has applied aspects of these approaches to explain how transfer operates in informal education, such as educational television, as well.

Particularly relevant to our research is the principle, adopted by several existing theories, that transfer can be elicited through *varied practice* (i.e., providing learners with multiple examples of the same concept or repeated practice of a skill in multiple contexts). Varied practice helps learners create a generalized mental representation of the material that is less context dependent, and more easily applied to new tasks and situations (e.g., Gick & Holyoak, 1983; Salomon & Perkins, 1989; Singley & Anderson, 1989). For informal education, we hypothesize that encountering similar content (e.g., a mathematical concept or problem-solving heuristic) in multiple contexts and media would lead, not only to a better grasp of the content, but also a greater likelihood of transfer to new problems as well.

Apart from the potential for varied practice to contribute to learning, research suggests that repeated, varied engagement with mathematical content can also promote positive attitudes toward mathematics (e.g., interest, motivation). Several theoretical approaches argue that interest in an academic subject develops from repeated, positive engagement with its content – for example, from repeated practice that results in positive emotional outcomes, from seeing the broader applicability and usefulness of the content, or from internalizing interest via encountering the content in engaging situations (e.g., Bransford, Brown, & Cocking, 1999; Hoffman, Krapp, & Renninger, 1998). Indeed, even the literature on attitude change in the context of advertising indicates that positive attitudes are more likely to occur from repeated exposure to persuasive messages – particularly if the precise content of the messages is somewhat varied (e.g., Kunkel, 2001; Petty, Priester, & Briñol, 2002). Under Hidi and Renninger's (2006) four-phase model of interest development, interest in an academic subject such as mathematics originates as interest sparked by the context in which the math is embedded (in this case, a math-based educational television program or game), and can evolve over time into interest in the underlying mathematics itself. In the domain of informal education, then, we hypothesize that experiencing math content via engaging materials across multiple media platforms contributes to the development of greater interest and positive attitudes toward mathematics.

Cyberchase

To investigate the nature of cross-platform learning, this study focuses on one such example for informal mathematics education, *Cyberchase*. Produced by Thirteen/WNET, *Cyberchase* is a popular and successful multiple-media mathematics project, targeted at children aged 8 through 11, that includes the following components:

- *Television series.* The animated *Cyberchase* series airs daily on PBS Kids Go! *Cyberchase* features three diverse youngsters who are summoned into Cyberspace to foil the dastardly Hacker. Each half-hour episode sends the team on a mystery based on a mathematics concept. Through their adventures, the series models mathematical reasoning, problem solving, and positive attitudes toward mathematics. Its underlying themes are that mathematics is everywhere and is infinitely useful. Nearly three million viewers – 43% of them African-American or Hispanic – tune in each week.

- *Web site. Cyberchase Online (www.pbskids.org/cyberchase)* complements the series with interactive games and puzzles, based on the same mathematical content as the television series. The site receives 1.8 million visits from nearly 700,000 unique visitors per month.

- *Outreach materials. Cyberchase* outreach engages children in hands-on math activities based on the television programs. Facilitators use a wide inventory of print and multimedia materials, including activity guides with DVDs of related episodes, for use in programs with children and workshops with colleagues. Partnerships with organizations such as Girls Inc. and the National Society of Professional Engineers bring *Cyberchase* outreach activities to communities nationwide, with a focus on girls and under-served children.

In addition to the above components (which serve as the focus of the present research), other materials include activity-based print materials, such as the *Cyberchase Activity Book* for kids and families. A traveling *Cyberchase* museum exhibit (produced in partnership with the Children's Museum of Houston) provides hands-on mathematics fun for millions of children and families alike.

This wealth of resources makes *Cyberchase* well-suited to the present research, for several reasons:

- *Cyberchase* offers a rich and varied library of existing materials in several media, which provides a powerful resource for both the naturalistic and experimental phases of the research.

- *Cyberchase* materials are designed to be complementary across media. For example, the same mathematics content is often addressed in an episode of the television series, an online game, and a hands-on outreach activity.

- Past research on the *Cyberchase* television series has proven its effectiveness as an educational tool. Sustained viewing has been found to result in statistically significant gains in children's mathematical problem solving and attitudes toward mathematics (Fisch, 2003; Rockman Et Al., 2002). The documented impact of the television series provides a useful baseline against which to assess outcomes resulting from use of multiple media.

- *Cyberchase* materials are widely viewed and used throughout the United States. As a result, the results of the proposed research pertain directly to real-world practice – they are more applicable and understandable than they might be if materials were created exclusively for the purposes of the research.

For all of these reasons, we have chosen to use *Cyberchase* materials in the proposed research. However, it is important to remember that, because so little published research literature exists regarding both informal mathematics education and children's learning from multiple media, the present study not only advances our understanding of the impact of *Cyberchase* itself, but also provides a first step toward understanding children's STEM learning from multiple media in general.

DESIGN AND METHOD

Sample

Participants were 672 children in nine public elementary schools in Michigan and Indiana. All of the children transitioned from third to fourth grade during the naturalistic phase, and remained in fourth grade during the experimental phase. Essentially the same sample of children participated in both the naturalistic and experimental phases of the research (described below), apart from some natural attrition as children moved into or out of their communities between third and fourth grade.

The sample was fairly evenly divided in terms of gender (52% girls, 48% boys), mathematics ability (31% high, 42% medium, 27% low), and whether math had been their favorite school subject prior to the study (43% yes, 57% no). Approximately 30% of the sample was comprised of minority children (17% African-American, 6% Latino, 4% Asian, 3% other).

Several schools were not willing to release socio-economic (SES) information regarding individual children. However, we were able to obtain aggregated statistics regarding eligibility for free or reduced lunch for each school as a whole. On average, 35% of the students at the participating schools were eligible for free or reduced lunch.

At the beginning of the study, we asked children how often they had viewed *Cyberchase* on television, and how often they had visited its Web site: never (defined as 0 times), a few times (1-5 times), a lot (6-10 times), or a whole lot (more than 10 times). For comparison, we also asked the same questions about two highly popular entertainment programs (*SpongeBob Squarepants* and *Scooby-Doo*), another educational series (*Liberty's Kids*), and their related Web sites. The following pair of tables summarizes their responses:

How often viewed on television	**Never (0 times)**	**A few times (1-5)**	**A lot (6-10)**	**A whole lot (more than 10)**
SpongeBob	2%	8%	9%	81%
Scooby-Doo	4%	26%	19%	51%
Cyberchase	38%	31%	12%	19%
Liberty's Kids	89%	7%	1%	3%

How often visited Web site	**Never (0 times)**	**A few times (1-5)**	**A lot (6-10)**	**A whole lot (more than 10)**
Nickelodeon (includes *SpongeBob Squarepants*)	23%	24%	15%	38%
Cartoon Network (includes *Scooby-Doo*)	38%	24%	11%	27%
Cyberchase	75%	15%	3%	7%
Liberty's Kids	98%	2%	0%	1%

As these tables show – and as expected – children's past viewing of top-rated entertainment programs (*SpongeBob Squarepants* and *Scooby-Doo*) dwarfed their prior exposure to the two educational programs (*Cyberchase* and *Liberty's Kids*). The vast majority of children had watched both entertainment programs more than a few times (90% for *SpongeBob Squarepants* and 70% for *Scooby-Doo*) – more than twice as many as had seen *Cyberchase* more than a few times.

In all four cases, children reported that they had watched each television series more often than they had visited its related Web site. This is consistent with prior research on *Cyberchase* media use (e.g., Fisch, 2005), as well as more general research on children's media use, which has found children to spend far more time watching television than going online (e.g., Rideout, Foehr, & Roberts, 2010).

Given the children's limited experience with *Cyberchase*, we expected that their prior exposure would not pose a confound for the present research. In fact, statistical analysis subsequently confirmed that there was no confound, as will be discussed in the Results section.

Research Design

One key consideration in designing any research study is whether to adopt a naturalistic approach, which allows findings to be generalized easily to the real world, or an experimental approach that provides a greater degree of control and allows researchers to attribute causality among the variables measured. To gain the advantages of both approaches, the present research was comprised of two phases. First, we conducted a *naturalistic phase*, whose primary purpose was to gauge children's naturalistic use of various *Cyberchase* media, and ways in which use of one medium feeds into another. This phase was followed by an *experimental phase*, in which treatment groups were exposed to various combinations of *Cyberchase* media. In this way, the design of the proposed research combined the power of experimental methods to yield clear, unambiguous evidence of effects with the real-world "meaningfulness" of naturalistic data.

The following table summarizes the schedule under which the study was run.

Schedule	Activities
April-June (start and end dates staggered to fit school schedules)	Background measures Assessment: Naturalistic phase pretest Naturalistic phase, part 1: approximately 6 weeks
October-December (start and end dates staggered to fit school schedules)	Naturalistic phase, part 2: approximately 6 weeks Assessment: Naturalistic phase posttest/Experimental phase pretest
January-March (start and end dates staggered to fit school schedules)	Experimental phase: 8-week treatment period Assessment: Experimental phase posttest *Cyberchase* interviews with teachers and children

At first glance, one might wonder whether use of *Cyberchase* during the naturalistic phase would pose a confound for the experimental phase. This is not the case, because the very nature of the naturalistic phase meant that children used (or did not use) *Cyberchase* media just as they would have in the absence of any research. Indeed, because we had a detailed record of use before the experimental treatment, we were able to run statistical analyses of experimental data that controlled for prior use – a level of control that would not have been possible if the children had not participated in the naturalistic phase.

Naturalistic Phase

As noted above, the primary goal of the naturalistic phase was to monitor naturalistic use of the *Cyberchase* television series and Web site over time. On one level, this was necessary simply to determine whether children actually do engage with an informal education project such as *Cyberchase* across several media platforms, or whether some children choose to watch the television series while others use the Web site (in which case the question of synergy across media platforms would be interesting but have little real-world relevance). On a second level, assuming that children make use of multiple media, naturalistic data were needed to explore how use of one medium for informal mathematics education might feed into use of other media (with implications regarding both synergy among media effects and children's motivation to engage in informal educational activities). Third, naturalistic data could help us identify predictors of children's use of *Cyberchase*, such as demographic factors (e.g., does use of *Cyberchase* differ as a function of children's pre-existing math ability?) and how use of *Cyberchase* fits into their overall pattern of media use. Finally, although the nature of the experimental phase was better suited to investigating children's learning from *Cyberchase*, we were also interested in whether naturalistic use was associated with differences in children's problem solving performance or attitudes toward mathematics. (Note, however, that we did not find enough naturalistic use to address the latter issue, but significant effects did emerge as a result of more extensive exposure during the experimental phase.)

The naturalistic phase tracked children's use of *Cyberchase* media over a 12-week period, approximately six weeks in the spring (when participants were in third grade) and six weeks in the fall (when they were in fourth grade). During this time, children were free to use (or not use) any *Cyberchase* media at home as usual. Once each week, they were asked to record any use of *Cyberchase* materials in a weekly "*Cyberchase* journal."

Measures. Measures administered during the naturalistic phase included:

- *Background questionnaire:* Administered at the beginning of the naturalistic phase, this measure gathered data on children's demographics (e.g., age, gender), prior viewing of *Cyberchase*, *Liberty's Kids*, *Scooby-Doo*, and *SpongeBob Squarepants*, as well as prior use of their related Web sites. A parallel measure was administered at the beginning of the fall data collection, to gather information about their use of *Cyberchase* and *SpongeBob Squarepants* during the summer.

- *Cyberchase journal:* Administered once each week, children used the journal to record their self-selected, voluntary use of the *Cyberchase* television series and Web site: whether (and on which days) they used each, how much time they spent on the Web site, and what activities they did on the Web site. For comparison, they answered the same questions about *SpongeBob Squarepants* and its Web site.

Samples of these measures can be found in Appendix A.

In addition, several pencil-and-paper measures of mathematical problem solving (pretest-posttest) and attitudes toward mathematics (posttest only) were also administered. These are described in detail in the discussion of the experimental phase below.

Apart from the information obtained directly from children via the above measures, we also collected teacher ratings of each child's mathematics ability (high/medium/low) and information regarding ethnicity and SES through the participating schools.

Experimental Phase

Whereas the purpose of the naturalistic phase was primarily to explore and document patterns of media use regarding *Cyberchase*, the experimental phase was focused on assessing impact regarding cross-platform learning. The experimental phase had three primary goals:

- To establish causality in studying the impact of multiple media on children's mathematical problem solving and attitudes toward mathematics.
- To assess how effects might differ as a function of children's exposure to different combinations of media components.
- To determine the respective contributions of each media component, and how these contributions interact to yield cumulative effects.

To that end, the experimental phase employed an experimental/control, pretest/posttest design that allowed us to investigate the impact of various combinations of *Cyberchase* media on the growth of children's mathematical problem solving (and, secondarily, their attitudes toward mathematics).

Experimental treatment. Over the course of an eight-week treatment period, intact classrooms of children were assigned to one of the following five experimental groups[1]:

[1] The constraints of daily classroom schedules required us to assign intact classrooms, rather than individual children, to each experimental group. However, to ensure that no systematic differences between classrooms would be confounded with the experimental treatment, multiple classrooms were assigned to each treatment group, and all of the treatment groups were roughly equivalent in their representation across age, gender, ethnicity, and mathematics ability. Indeed, subsequent statistical analysis confirmed that there was no significant

- *DVD Only group:* Each week, children were shown three half-hour episodes of *Cyberchase* in school (a total of 24 episodes).

- *Web Only group:* Each week, children played a mathematics-based game on the *Cyberchase* Web site (a total of 8 games), but were not shown the TV series.

- *DVD + Web group:* Children were shown three episodes of *Cyberchase* per week. Once each week, they also played an online game whose mathematical content was (in most cases) aligned with at least one of the TV episodes they viewed.

- *All Materials group:* Children followed the same schedule as in the previous group. Once each week, they also engaged in a hands-on *Cyberchase* outreach activity that involved the same mathematical content as in one or more TV episodes (a total of 8 hands-on activities).

- *No Exposure (i.e., control) group:* Children were not exposed to any of the above materials. Instead, each week, they were shown three half-hour episodes of an age-appropriate series about American history (*Liberty's Kids*).

Exposure to the various *Cyberchase* media was designed to emulate real-world use of these materials. Past surveys (Fisch, 2005, 2006) and Nielsen ratings data indicate that, of the *Cyberchase* media components that currently exist, the television series has the greatest reach, followed by the Web site. (The greater use of television over Web was also evident in data from the naturalistic phase of the present study, as reported in the Results section below.) Thus, most of the treatment groups were exposed to the television series and/or Web site. Similarly, each group's frequency of exposure was informed by data on frequency of real-world use, as well as the maximum that schools and after-school programs were able to accommodate.

It is worth noting that, in designing the treatment, we recognized that children would spend more time with some forms of *Cyberchase* media (e.g., television) than others (Web site or hands-on activities), which could contribute toward some media producing greater effects than others. However, the alternative would have been to have children spend equal time with each media component, which is not representative of naturalistic use; the artificiality of such a treatment would have limited the generalizability of our results to the real world. For this reason, and because the purpose of the research was to investigate synergy among media rather than to attempt to determine which medium is "best" as a tool for informal education, we decided to align the treatment with real-world use so as to maximize the generalizability of the data.

To select specific television episodes, online games, and hands-on games for the treatment, we reviewed the available *Cyberchase* library, and chose materials that fell within two broad areas of mathematics content: organizing data (e.g., graphs, combinatorics, predicting from data) and measurement (e.g., size and scale, elapsed time, proportional reasoning). In some weeks of the

---

effect of teacher or classroom on any of our measures (apart from treatment effects), as will be discussed in the Results section.

treatment, mathematics content was closely aligned across media; for example, one television episode, one online game, and one hands-on activity all concerned "body math" – that is, proportional relationships among body parts (e.g., a person's foot is approximately the same length as his/her forearm). In other weeks, the mathematical topic was not closely aligned across media, although the same sorts of problem-solving strategies and heuristics could be applied in all three. A complete list of all of the materials in the treatment, along with the schedule under which they were administered, is presented in Appendix B.

Children in the No Exposure group were not presented with any of the *Cyberchase* materials. Instead, they watched three episodes per week of the television series *Liberty's Kids*. Like *Cyberchase*, *Liberty's Kids* is an animated, educational series that is aimed at approximately the same age group. However, whereas the educational content of *Cyberchase* focuses on mathematics, *Liberty's Kids* deals with American History at the time of the Revolutionary War.

Measures. The following measures were administered during the experimental phase:

- Hands-on mathematical problem-solving tasks, with essentially isomorphic versions of each task administered in the pretest and posttest.
- Paper-and-pencil problem-solving tasks, administered in the pretest and posttest.
- Online tracking data that automatically recorded every click children made while playing three of the interactive games on the *Cyberchase* Web site. The tracking data lent insight into children's mathematical thinking while playing the games.
- Several paper-and-pencil measures of interest and confidence regarding various mathematical activities (administered in the pretest and posttest) and children's orientations toward pursuing challenges (posttest only).
- Finally, to help interpret these data, the posttest was followed by supplementary interviews with children (regarding their experience with *Cyberchase*) and teachers (to gather their perspective on children's learning from *Cyberchase*, as well as their own experiences with the materials).

Each of these measures is described in greater detail below.

Measures – problem solving. Several past studies of children's mathematics learning from television have employed hands-on problem-solving tasks in the context of task-based interviews (e.g., Fisch, 2003; Hall, Esty, & Fisch, 1990). The present study, too, included hands-on tasks, but the design of our problem-solving assessments was informed largely by the *thought-revealing* or *model-eliciting activities* approach described in the mathematics education literature (e.g., Lesh & Doerr, 2002; Lesh, Hamilton & Kaput, 2007; Lesh et al., 2000). In essence, thought-revealing activities are a means of blending instruction and assessment, by presenting students with rich, meaningful problems that can be approached in a variety of ways (including both mathematical and non-mathematical approaches). Students work on each problem in groups, with the goal of not only solving the problem, but also describing a more generalized procedure by which other, similar problems can also be solved. Each problem is designed to yield insight into the students' mental models of its mathematics content, and how these models evolve; typically, students' mental models become more sophisticated and accurate as the group continues to work on the problem over an extended period of time (sometimes several days).

To serve as a useful tool for a pretest-posttest study, as opposed to classroom instruction, the thought-revealing activities approach had to be adapted in several ways. First, and perhaps most important, thought-revealing activities for the classroom are designed to serve as an instructional tool as well as an assessment. Yet, for the purposes of a research study, the educational value of the assessment tasks could not be so strong as to overwhelm any effects of the experimental treatment (although they still needed to be rich enough to lend insight into the process of children's problem solving). To that end, we limited both the scope of each task and the amount of time that children were given to work on it. Second, because of these imposed limits, we were unsure whether children would be able to not only solve the problem but also reflect on and abstract their solution into a more generalized procedure for solving other problems in the future. Thus, apart from asking children for their solutions, we also asked them two levels of questions for each problem: First, we asked them to recount the process they used to solve this particular problem, and then we asked them to describe a process that could be used to solve similar problems in the future. Third, for the purposes of an experimental study, we had to create coding schemes that operationalized aspects of children's process and solutions in standardized ways that would be reliable from pretest to posttest and across experimental groups. Finally, in addition to hands-on tasks, we also devised paper-and-pencil problem-solving tasks (and related coding schemes) that captured the same spirit as the hands-on tasks, but could be completed on paper by children working alone.

Children completed paper-and-pencil assessments of problem solving at three points: at the beginning of the naturalistic phase, at the end of the naturalistic phase, and at the end of the experimental phase. Thus, the naturalistic and experimental phases were each framed by a pretest and posttest, with the second wave of assessment serving simultaneously as both the posttest for the naturalistic phase and the pretest for the experimental phase.

During each wave of assessment, children were given two paper-and-pencil tasks, one of which was a measurement task, and the other involved organizing data. Each measurement task presented children with a figure that included four zigzagging paths; children were asked to figure out which of the presented paths was longest or shortest. Children were awarded points based on their use of measurement and whether they indeed chose the longest/shortest path. Each organizing data task presented children with a table of data (e.g., the amount of food collected in response to various methods of publicizing a food drive for needy families), and asked them to interpret the data in the table and make predictions about what might happen in the future given current trends. Children were awarded points based on their ability to interpret the table, and to use data to make and justify well-founded predictions. Pretest and posttest tasks were essentially isomorphic to each other, employing the same underlying mathematics in a different surface context. Samples of the paper-and-pencil tasks, and their coding schemes, can be found in Appendix C.

Hands-on tasks were administered at the beginning (pretest) and end (posttest) of the experimental phase. As in the paper-and-pencil assessments, children completed two hands-on tasks each time, a "body math" task that required measurement and proportional reasoning, and an organizing data task that involved combinatorics. Whereas the paper-and-pencil tasks were completed by each child individually, however, children worked on the hands-on tasks in groups of three. Researchers observed each group of children as they worked on the task, and interviewed them afterward about their process and solution.

For example, the pretest body math task cast children in the role of detectives who had to figure out as much as possible about the perpetrator of a mysterious crime.

[A corner of the room is set up as follows, to simulate a crime scene (note: the measurements are for reference; they are not be told to children, although children can choose to measure the objects themselves):

- *A 5-foot-wide frame sits on the floor in front of the wall. Torn paper hangs around the inside edge of the frame, to simulate a picture that's been cut out. There are matching dirty handprints on the left and right edges of the frame where someone held it.*
- *A pair of 10-inch-long footprints face the wall.*
- *A hat (20" around) lies nearby, with one brown hair (represented by yarn) inside]*

In this puzzle, we're going to pretend that you are detectives. Famous detectives like Sherlock Holmes can tell a lot about crooks from the clues they leave behind. For example, they can use footprints to figure out how tall someone is, how much they weigh, whether they walk with a limp, and so on. Fingerprints can tell them even more, because everybody's fingerprints are different. But in this crime, the crook didn't leave any fingerprints behind, although there are some other kinds of clues you can use. Ready?

Here's what happened: Imagine this is a museum where a priceless painting was stolen. *[While demonstrating:]* The thief came in at night, stood here *[in footprints]*, and took the painting off the wall *[one hand on either side of frame]*, leaving these dirty handprints on the frame. Then, the thief cut out the painting, rolled it up, and ran away. But during the getaway, the thief's hat fell off. It's lying right here, with a blonde hair inside.

Your job is to use these clues to figure out as much as you can about what the thief looks like: how tall the thief is, what color hair the thief has, and so on. You might be able to tell a lot about the thief, or maybe just a little. Either way, each thing you figure out will help narrow down the search, and that'll make it easier to catch the crook. So each piece of information is important.

To help you, you can use anything you want from this kit of detective tools – or you don't have to use any of the tools at all. *[See Appendix F for list of materials in kit.]*

Okay, now I'm going to give you a little time to figure out the puzzle. You can do whatever you want to help you figure it out, and if you want to use any of this stuff [kit], you can. I'll be over there working if you need me. Otherwise, when you're ready, you can call me and we'll talk about what you think the thief looks like. Any questions? Okay, let's begin.

(continued on next page)

(Detective task continued)

PART 2

Good news: Thanks to your description, the police caught the thief and got the painting back! The Chief of Police is so happy that she wants you to teach all the other detectives how to catch crooks just like you did.

So now, I want you to think about how you'd teach someone <u>else</u> to use these kinds of clues to figure out what a crook looks like. Remember, the Police Chief doesn't want you to come up with a description of <u>one</u> crook this time. Instead, the Chief wants you to give the detectives some directions for <u>how</u> to figure it out, so they can use the same steps every time they have to solve this kind of crime.

Take a little time to think about it. When you're ready, let me know and we'll talk about what you're thinking. Any questions? Okay, let's begin.

This mystery could be approached in a variety of ways, both mathematical and non-mathematical (e.g., using the brown "hair" as evidence that the thief had brown hair). A mathematically complete and sophisticated answer, however, required children to apply proportional reasoning to use presented clues (e.g., the size of the footprints and/or distance between handprints) and thus draw inferences about the thief's height and the size of various parts of his or her body. Similarly, the posttest task cast children as sculptors in a wax museum whose task was to figure out dimensions for a life-size statue of basketball player Shaquille O'Neal, based on a photo and an outline of his footprint.

Organizing data tasks required children to construct a schedule for a series of ping pong or soccer matches, providing for all possible combinations of competitors while also meeting several constraints for scheduling. As in the body math tasks, this pair of tasks could be approached in a variety of ways, such as using multiplication for combinatorics, or physically manipulating cards with competitors' names to create possible matches.

Children's performance in each hands-on task was coded in two ways, one reflecting the process they used while working on the task, and the other representing the mathematical sophistication of the group's solution. For process, a detailed coding scheme was devised to identify the strategies and heuristics that children used while working on the task (e.g., standard or nonstandard measurement, looking for patterns, trial and error). This scheme focused on the following strategies and heuristics.

Heuristics Coded in Process Score

- Recall information
- Gather information
- Measure: Ruler
- Measure: Nonstandard manipulatives
- Estimate, approximate
- Calculate
- Manipulate: Use objects
- Manipulate: Change objects
- Trial & error, guess & check
- Write: List, table, chart
- Write: Picture, diagram
- Write: Other
- Transform problem
- Look for patterns
- Reapproach problem
- Reasonableness
- Alternate ways to solve
- Related problems

The coding scheme for solution scores identified several levels of sophistication, based on the key mathematical concepts underlying the task and the systematicity of the solution. Because children worked on the task in groups of three, one process score and one solution score was assigned to each triad as a whole.

Coding for the paper-and-pencil tasks was similar, with two primary differences. Because researchers did not observe children as they worked on the paper-and-pencil tasks, children were assigned solution scores but not process scores. Also, because children completed these tasks individually, every child received a score for each paper-and-pencil task (whereas scores for hands-on tasks were assigned to groups rather than individual children).

Samples of the hands-on tasks, along with the coding schemes for each task's solution score, can be found in Appendix D. The coding scheme for process scores can be found in Appendix E. Note that the same coding scheme was used to assign process scores for all of the hands-on tasks.

Measures – online tracking data. In keeping with current trends in mathematics education toward blending instruction and assessment (e.g., Kelly & Lesh, 2000; NCTM, 1993), several *Cyberchase* online games served simultaneously as both instruction (i.e. part of the experimental treatment) and assessment (i.e. a means of gauging children's mathematical problem solving). Three games were chosen to serve in this capacity: Railroad Repair, a game about adding decimals to create given sums (*http://pbskids.org/cyberchase/games/decimals/*); Sleuths on the

Loose, a "body math" game about proportional reasoning (*http://pbskids.org/cyberchase/games/bodymath/*); and Pour to Score, a game about measuring and creating given quantities of liquid (*http://pbskids.org/cyberchase/games/hardproblems/*). Custom-built tracking software was created to record each player's mouse clicks and keyboard input automatically as he or she played the game. Parallel to the hands-on assessments discussed above, detailed coding schemes were created to draw on characteristic patterns of responses and errors to identify the sophistication of the player's strategies for playing the game, as well as the number of items he or she answered correctly.

For further detail on this innovative methodology, and its potential for assessing mathematical problem solving, please see Fisch et al. (in press). Sample coding schemes can be found in Appendix G of the present report.

Measures – attitude. In approaching attitudes toward mathematics, we faced a significant challenge, because existing measures of mathematics attitude (e.g., Fennema & Sherman, 1976) are of limited usefulness in the context of informal education, for several reasons. First, their items typically focus heavily on in-school mathematics, whereas the focus of informal education is outside of school. Second, they typically make extensive use of the word "math," which children (and adults) often interpret as meaning little more than numbers and arithmetic, rather than the far broader range of subjects that actually comprise mathematics (e.g., Debold et al., 1990; Kulm, 1980).

For these reasons, we created new measures to assess attitude both in and outside the context of school. Our assessments were informed by Hannula's (2002) recent approach to assessing attitude in multiple contexts, as well as the research literature on interest and motivation (e.g., Dweck & Elliott, 1983; Hidi & Renninger, 2006; Hoffman, Krapp, & Renninger, 1998) and measures used in past summative studies of the impact of informal mathematics and science education (e.g., Fay et al., 1995; Debold et al., 1990; Nicholson, Hamm, & Weiss, 1991). These measures focused on three dimensions of attitude toward mathematics: interest, confidence, and motivation.

One paper-and-pencil measure (administered in the experimental pretest and posttest) was a set of interest and confidence scales, based on the scale used in Fay et al.'s (1995) summative evaluation of the science and technology television series, *Cro*. To avoid ambiguous uses of the word "math," the scale instead asked children to rate their interest and confidence in a variety of specific activities. Four categories of activities were included:

- *Cyberchase items*: Activities that corresponded directly to mathematics-based activities from the *Cyberchase* materials included in the treatment
- *Non-Cyberchase items*: Mathematics-based activities that were not included in the *Cyberchase* materials in the treatment (although they are the subjects of other *Cyberchase* materials that were not used in the study)
- *School math items*: Activities typical of mathematics in school
- *Non-math items*: Activities that were less inherently mathematics-based

For example, some items in this measure asked children to rate their interest and confidence in figuring out: ways to keep track of time without using a clock (*Cyberchase* item), whether a game of chance is fair for all of the players, (non-*Cyberchase* item), the best way to study for a math test (school math item), and the history of their home town (non-math item). Interest in each activity was rated on a five-point scale of *very interesting, a little interesting, so-so, a little boring*, or *very boring*. Confidence in one's ability to do the activity was rated on a five-point scale of *definitely could do it, maybe could do it, not sure, maybe could* not *do it,* and *definitely could* not *do it*.

A second paper-and-pencil measure (administered only in the experimental posttest) grew out of past literature that has examined motivation in terms of children's goals and orientations toward challenge – specifically, whether their motivation is intrinsic to the task itself (often referred to as *learning goals* or *mastery goals*) or stems from extrinsic factors such as external praise or rewards (often referred to as *performance goals*; see, e.g., Pintrich & Schunk, 2002). In this measure, children were presented with three stories. Each story concerned three children who encounter some difficulty while working on a difficult problem; one of the children suggests that they continue to work because the task is interesting (consistent with learning/mastery goals), a second suggests continuing simply because they started the task and so should finish it (consistent with performance goals), and a third suggests asking someone else to tell them the right answer (performance goals). In keeping with Hannula's (2002) recommendation of considering context while measuring attitudes toward mathematics, the problem in each story was set in different context: solving a classroom math problem, attempting to complete a level in a video game, and making a gift for a parent. Children were asked to indicate the characters in each story who sounded most and least like themselves, and to explain why. Samples of these paper-and-pencil measures can be found in Appendix H.

Finally, apart from its role as a problem-solving assessment, we also drew on the online tracking data discussed above to provide a behavioral measure of motivation. Typically, it is difficult to use variables such as time spent in a mathematical task as a measure of motivation, because skill and motivation are confounded; if a child spends only a brief period of time in the task, is it because he or she is unmotivated, or highly skilled and able to complete the task quickly? In our online tracking data, however, pilot data revealed that some players not only completed an entire game, but also continued to play again, beyond the end of the game. By restricting our motivation analysis to only those children who reached the end of the game, we were able to control for ability (because all of these children had sufficient ability to reach the end of the game); continuing to play beyond the end of the game was a clear indicator of greater motivation than simply stopping when the player reached the end. Thus, by comparing children in the Web Only group to children who used multiple *Cyberchase* media, we could determine whether the use of multiple media increased children's motivation to engage further in mathematics-based games.

Reliability and validity. Although all of our measures were grounded in past literature on research and math assessment, all of them were created for the purposes of this research. Thus, prior to the study, we conducted an extensive pilot phase to test the reliability and validity of the measures, discard any measures that did not perform adequately, and refine the design of the measures that remained. During this time, the research design and measures were also reviewed and approved by the institutional review board at Michigan State University, to ensure compliance with standards for the ethical treatment of human subjects.

As reported elsewhere in greater detail (Fisch, 2007), the measures were found to be sufficiently valid and reliable. Interrater reliability exceeded $r = .90$ ($p < .01$) for paper-and-pencil problem-solving tasks and $r = .80$ ($p < .01$) for solution scores to hands-on tasks, and there was 79% agreement regarding the problem-solving heuristics used in hands-on tasks (Cronbach's alpha = .69). Performance on both types of tasks was significantly correlated with both teacher ratings of math ability and performance on a scale comprised of items taken from the mathematics portion of the fourth-grade National Assessment of Educational Progress (NAEP), although a few of these correlations narrowly missed attaining significance due to the small sample size used in the pilot test.

Principal component analyses suggested that the subscales of the attitude scales were well defined, in that virtually all of the mathematics-based items (*Cyberchase*, non-*Cyberchase*, and school math) clustered around one factor, whereas the non-math item *Figuring out how to take care of a pet* did not. Interrater reliability for the attitude scales was quite high: $r = 1.00$ ($p < .01$) for the interest scale, and $r = .996$ ($p < .01$) for the confidence scale.

Supplementary interviews. To aid in interpreting the data from the above measures, the posttest was followed by supplementary interviews conducted individually with participating children and teachers. Child interviews focused on children's experience with *Cyberchase* during the treatment, any follow-up activities or discussion that might have occurred at home, and children's perceptions of *Cyberchase*'s appeal and embedded problem solving, as well as their perceptions of their own learning (if any). Because this interview concerned *Cyberchase* itself, it could only be administered to those experimental groups that used *Cyberchase* materials (i.e. not the No Exposure group).

Teacher interviews gathered information on topics such as the teachers' perceptions of their students' learning from *Cyberchase*, how *Cyberchase* was integrated into their classrooms, and their perceptions of *Cyberchase*'s educational value and usefulness in the classroom.

Interview protocols for the child and teacher interviews are presented in Appendix I.

RESULTS

Patterns of Use of Multiple Media (Naturalistic Phase)

Overview of Results: Naturalistic Use

Use of each form of *Cyberchase* media (TV and Web site) was fairly consistent over time. Those children who watched *Cyberchase* on TV in one month tended to do so in subsequent months as well, and a similar pattern was found for month-to-month use of the Web site.

Children's use of *Cyberchase* also tended to span media; each month, children who watched the TV series more frequently also tended to visit the Web site more often. Thus, in naturalistic use, some children do indeed use multiple media when they are available (which lends real-world validity to the question of how children learn from multiple media).

Because most users' first encounter with *Cyberchase* occurred long before we began collecting data, the present data cannot determine which medium came first. However, past research found that children more often begin by watching the TV series and subsequently expand to using the Web site as well.

Overall Amount of Use

Earlier in this report, we reported the degree to which children said they had watched the television series or visited the Web site prior to participating in the study (see the description of the sample on page 10). Children's use of *Cyberchase* media during the naturalistic phase was fairly consistent with their earlier self-reports of prior use. The following two tables draw on data from children's weekly *Cyberchase* journals, in which they recorded the number of times they watched *Cyberchase* or *SpongeBob Squarepants* on television, or visited the related Web sites during the naturalistic phase. For the sake of comparison, we have recoded the data into the same categories used on page 10.

Number of children who watched on TV during the naturalistic phase	0 times	1-5 times	6-10 times	More than 10 times
Cyberchase	30%	42%	14%	14%
SpongeBob	10%	10%	11%	69%

Number of children who visited Web site during the naturalistic phase	0 times	1-5 times	6-10 times	More than 10 times
Cyberchase	35%	47%	11%	7%
SpongeBob	54%	21%	9%	16%

Several points regarding these data are noteworthy:

Children's overall amount of use of *Cyberchase* media during the naturalistic phase was fairly consistent with their earlier self-reports of use prior to the study. This consistency lends validity to the data from both measures, and suggests that their level of *Cyberchase* media use during the naturalistic phase was indeed naturalistic; it does not appear to have been overly inflated by the children's knowledge that they were participating in a research study.

During the naturalistic phase (as in their use prior to the study), children watched *SpongeBob Squarepants* more often than *Cyberchase* during the naturalistic phase. Again, this is consistent with Nielsen ratings that show *SpongeBob Squarepants* to be highly popular among this age group. Interestingly, use of the *SpongeBob* Web site was somewhat split; although more children visited the *SpongeBob* site frequently (more than 10 times) than *Cyberchase*, more children never visited it as well.

As mentioned earlier, the lower levels of use for *Cyberchase* held both advantages and disadvantages for the present study. On the one hand, there was not enough use to produce significant pretest-posttest effects on problem solving during the naturalistic phase (although such effects did arise as a result of more extensive use during the experimental phase). On the other hand, because children did not spend enough time with *Cyberchase* to produce educational effects during the naturalistic phase, the results of the experimental phase could be interpreted cleanly; there was no confounding effect stemming from children's earlier use during the naturalistic phase.

Still, although naturalistic use may not have been sufficiently prevalent to elicit change in problem-solving performance, it was sufficient to explore patterns of multiple-media use over time. Path analysis, via structural equation modeling, was conducted to address issues such as whether children's use of each *Cyberchase* medium was consistent over time, whether use of one form of *Cyberchase* media was associated with use of another, and whether use might be predicted by demographic variables or children's more general use of television or the Web.[2] We consider each of these issues in turn.

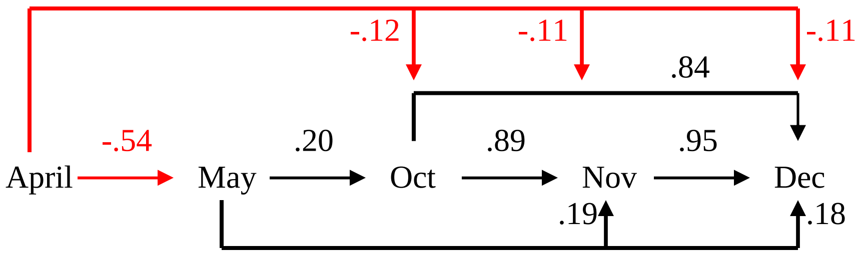
Consistency of Use Over Time

The following figures represent relationships among children's use of a given medium (e.g., watching the *Cyberchase* television series) from month to month during the naturalistic phase. All of the arrows and numbers shown are statistically significant at $p < .05$ or greater; missing arrows reflect relationships that were not statistically significant.

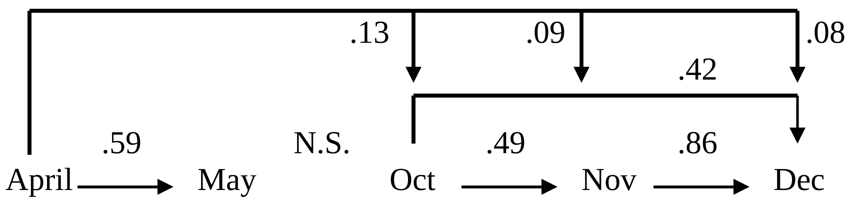
Please note that the numbers shown in these figures are estimates of the strength of each path, not correlation coefficients (and, thus, should not be interpreted as correlations). However, like correlations, these values indicate the strength of each relationship, and the degree to which one variable predicts another.

[2] Further detail on these and other statistical analyses presented in this report can be found in Appendix J.

Cyberchase TV

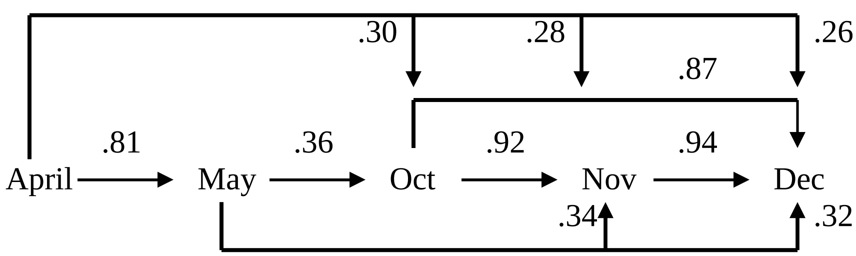


Cyberchase Web

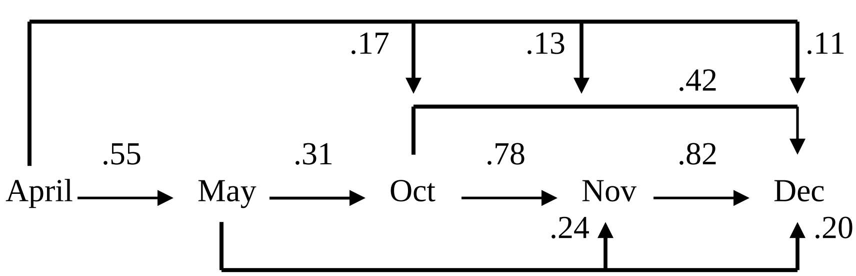


(May N.S. to Nov and Dec)

SpongeBob TV



SpongeBob Web



As these figures illustrate, use of each form of *Cyberchase* media (TV and Web site) was fairly consistent over time. Those children who chose to watch *Cyberchase* on TV in one month also tended to do so in subsequent months. A similar pattern was found for month-to-month use of the Web site, and use of *SpongeBob Squarepants* was found to be consistent over time as well. Thus, use of each *Cyberchase* medium typically was not a one-time experience. Rather, it tended to continue at a similar level over an extended period of time (regardless of whether a particular child chose to use *Cyberchase* many times each month or only a few times per month). As one might expect, relationships between months were generally strongest for neighboring months; for example, the path estimate between *Cyberchase* television viewing in May and December was .18, but the path estimate between November and December was .95 (although both estimates were statistically significant).

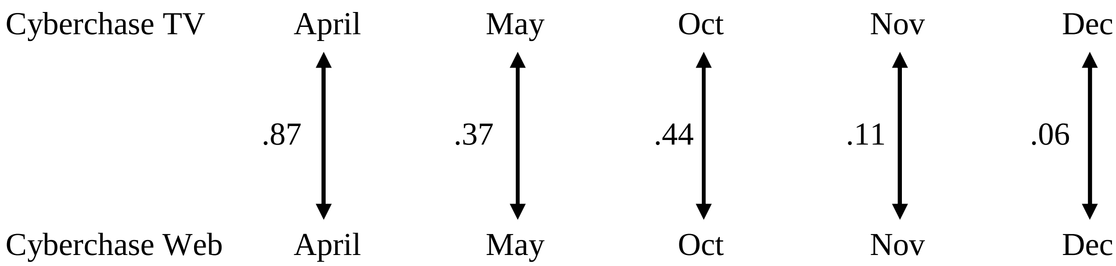
The one notable outlier to this trend was children's viewing of the *Cyberchase* television series in April. Unlike the other months (and other media), *Cyberchase* viewing in April was negatively related to the remaining months. Three possible explanations of this unexpected result seem likely: (a) It is possible that viewership of *Cyberchase* was inflated in April due to children's curiosity upon entering the study (whereas most children were already familiar with *SpongeBob Squarepants*, so no such effect appeared for the latter series). (b) Viewership in April may have been inflated due to Spring break, during which children had more free time available to watch *Cyberchase*. (c) Although children in all of the participating schools completed *Cyberchase* journals for 12 weeks during the naturalistic phase, the constraints of differing school schedules caused some schools to begin the journals sooner than others. Because only some of the children began filling out their journals in April, the restricted range of data may have affected the analysis for April (whereas May, by contrast, was positively related to all of the remaining months). We suspect that all three of these factors may have contributed to some degree.

Despite this one unexpected finding for television in April, however, the bulk of the data make it clear that the degree to which children use a given medium for informal education (like their use of a given medium for entertainment) tends to remain stable over a period of several months.

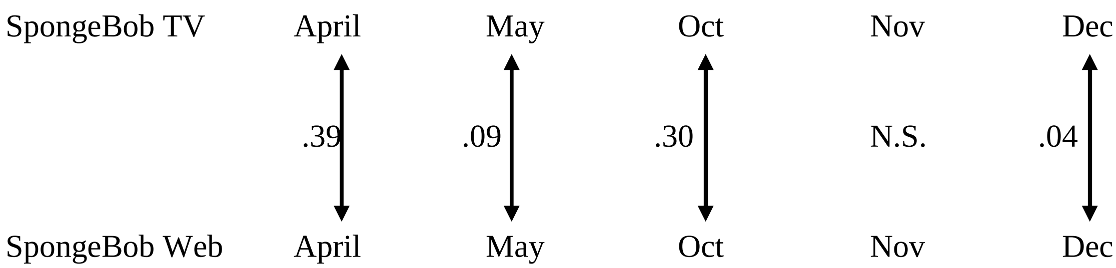
Cross-Platform Use: Television and Web

Just as use of a given medium was found to be relatively stable from month to month, children's use of *Cyberchase* also tended to carry over from one medium to another. The following pair of figures illustrates relationships between use of a given television series and related Web site within each month. As above, all of the arrows and path estimates shown in the figures are statistically significant at $p < .05$ or greater.

TV-Web relationship within months (*Cyberchase*)



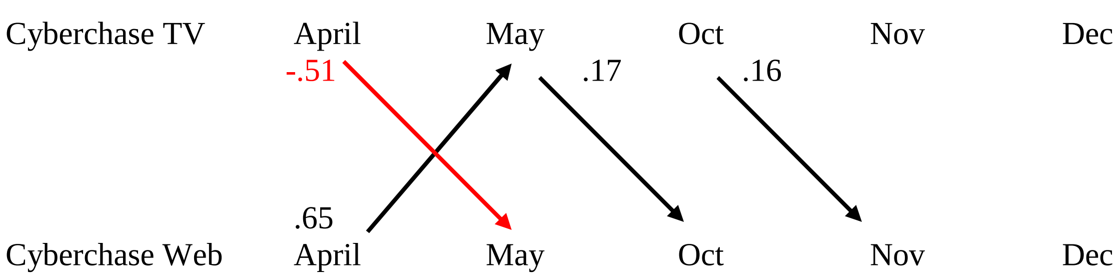
TV-Web relationship within months (*SpongeBob*)



As the above figures demonstrate, use of the *Cyberchase* television series and Web site were significantly related in every month. The same was true, in almost every month, for *SpongeBob Squarepants* as well. Thus, these data are consistent with one of the central assumptions underlying the present research: At least in the context of *Cyberchase*, children will engage with informal educational media across platforms when related multiple media are available.

With that in mind, a natural next question might be which medium "comes first" – whether use of television typically leads to use of a related Web site or vice versa. To explore this question, we conducted a set of path analyses that compared use of one medium in a given month to use of the other medium in the following month. The results of these analyses are summarized in the following pair of figures.

TV-Web relationship between months



TV-Web relationship between months

SpongeBob TV	April	May	Oct	Nov	Dec
		No significant relations found between months			
SpongeBob Web	April	May	Oct	Nov	Dec

As the above figures illustrate, the relationship between television and Web did not always proceed in the same direction. For *Cyberchase*, there were two instances in which the relationship proceeded from television to Web, and one in which it proceeded in the opposite direction. (In addition, as in the analysis of television viewing across months, television viewing in April was an outlier, in that it was negatively related to Web use in May.) For *SpongeBob Squarepants*, surprisingly, there were no instances in which use of one medium predicted use of the other medium in the following month, despite the fact that significant relationships between television and Web did emerge within months, as discussed earlier.

We believe that clearer directional trends were not found because the children who used either *Cyberchase* or *SpongeBob* during the naturalistic phase typically were not using it for the first time (as seen in our data on children's media use prior to the study). Because they already had some experience with one or both media platforms, the data from the naturalistic phase could not truly speak to the question of which platform "came first."

However, prior research does suggest an answer to this question in the case of *Cyberchase*. Nielsen ratings and online metrics indicate that both *Cyberchase* and *SpongeBob Squarepants* reach larger audiences on television than online (and American children spend considerably more time with television than the Web overall; Rideout et al, 2010), and a survey of parents of *Cyberchase* viewers found that children typically began watching *Cyberchase* on television at younger ages than they began using the Web site (Fisch, 2005). Certainly, the path between television and Web use can proceed in both directions -- and, indeed, the potential for multiple entry points is one of the strengths of multiple media. In most cases, though, it appears that children first encounter projects such as *Cyberchase* on television, and those children who find the material sufficiently appealing to become *Cyberchase* fans continue to engage with *Cyberchase* over time and across other media platforms.

Predictors of Use

Having established that use of *Cyberchase* media was consistent over time and spanned multiple media platforms, we next turn to the question of how use of these media might be predicted or influenced by use of other media and external demographic variables. To do so, we considered several indicators of children's use of *Cyberchase* media: their reports of *Cyberchase* television and Web use prior to the study, and their use during the naturalistic phase (as reflected in the number of times they used each medium and the amount of time spent). The following correlation matrix summarizes the degree to which each of these variables was predicted by children's gender, ethnicity (coded as minority vs. nonminority), mathematics ability, and whether they cited math as their favorite subject in school, as well as their use of *Cyberchase* media prior to the study.

	Cyberchase TV viewing (before study)	Cyberchase Web use (before study)	Number of times watched Cyberchase on TV (naturalistic phase)	Number of times visited Cyberchase Web site (naturalistic phase)	Total time spent: Cyberchase TV (naturalistic phase)	Total time spent: Cyberchase Web site (naturalistic phase)
Gender	0.04	0.09*	0.02	0.06	0.02	0.05
Ethnicity	-0.11	0.11	0.07	0.11	0.07	0.13*
Math ability	-0.03	0.02	-0.08	0.01	-0.07	0.00
Favorite subject math	-0.05	-0.03	0.07	0.11*	0.07	0.02
Cyberchase TV viewing (before study)	---	0.49***	0.27***	0.10*	0.26***	0.09*
Cyberchase Web use (before study)	0.49***	---	0.26**	0.25***	0.25***	0.32***

*Statistically significant, p < .05. ***Statistically significant, p < .001.

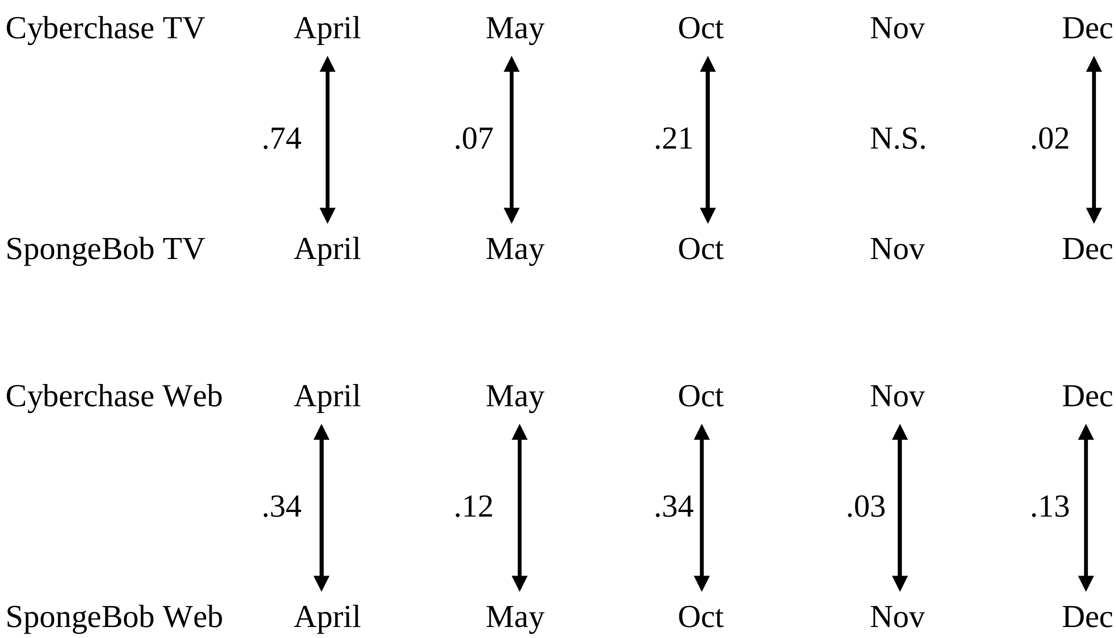
Data from Nielsen ratings and past studies have indicated that *Cyberchase* is successful in reaching both boys and girls and children of various ethnicities (e.g., Fisch, 2003, 2005). As this table shows, the present data confirm these findings. Neither gender nor ethnicity predicted the degree to which children chose to watch *Cyberchase* on television, suggesting once again that the series had roughly equivalent reach among a diverse audience. Only two correlations regarding the Web site reached significance, in that girls were more likely to have visited the site prior to the study ($r = .09$, $p < .05$), and minority children spent more time on the site during the naturalistic phase ($r = .13$, $p < .05$).

Similarly, past research has shown that children are typically attracted to *Cyberchase* for the first time because of its entertainment value rather than their desire to seek out a program about mathematics (Fisch, 2005). In the present data, too, neither mathematics ability nor naming mathematics as a favorite subject was a significant predictor of viewing the *Cyberchase* television series. Only one such correlation emerged as significant, in that children who named math as a favorite subject were significantly more likely to visit the *Cyberchase* Web site ($r = .11$, $p < .05$). Thus, *Cyberchase* appeared to have similar reach among both "math kids" and non-"math kids."

In contrast to the limited predictive power of the preceding demographic variables, the strongest and most consistent predictors of *Cyberchase* media use lay in children's prior use of *Cyberchase* media. Just as the naturalistic data found that self-selected use of *Cyberchase* was consistent over time and across platforms (as discussed earlier), the correlation matrix revealed that children's use of each *Cyberchase* medium (television and Web) prior to the study significantly predicted their use of both types of media during the naturalistic phase. This was true in terms of both the number of times children used the television series and Web site, and the amount of time they spent.

Given this relationship among use of *Cyberchase* media, it is also reasonable to ask whether use of each medium might be related to their use of television and the Web in general. To that end, structural equation modeling was used to conduct a path analysis to examine the relationship between *Cyberchase* and *SpongeBob Squarepants* within each month of the naturalistic phase. The following figures present the relationships found between the two television series and between the two Web sites.

Relationship between *Cyberchase* and *SpongeBob Squarepants*



As these figures illustrate, there were significant positive relationships between children's viewing of the two television series in almost every month, and significant positive relationships between the two Web sites in every month. Thus, the choice to use educational or entertainment media was not an "either-or" dichotomy. Rather, children who engaged more with one media genre were likely to engage more with the other genre as well.

In sum, children's use of *Cyberchase* media was predicted most consistently by their use of other media – their prior use of each type of *Cyberchase* media, and their use of entertainment media. Use of the *Cyberchase* television series was not predicted by either gender, ethnicity, mathematics ability, or naming math as a favorite subject, because *Cyberchase* reached all of these groups to a similar degree. However, girls, minority children, and children whose favorite subject was math all showed somewhat greater use of the *Cyberchase* Web site.

Educational Impact and Cross-Platform Learning

Having documented ways in which children used multiple *Cyberchase* media during the naturalistic phase, we now turn to experimental data that evaluated the educational benefits of children's engagement with such media. We will consider these effects within three broadly defined areas: impact on mathematical problem solving, cross-platform learning from multiple media, and effects on children's attitudes toward mathematics.

Overview of Results

Consistent with past research on the *Cyberchase* television series, use of *Cyberchase* resulted in significant gains on both the process of children's mathematical problem solving and the sophistication of their solutions. Interestingly, learning from *Cyberchase* was not manifest in children's simply doing a greater number of things while working on the tasks, but rather in their using a greater *variety* of strategies and heuristics, and in using those strategies and heuristics more effectively. As expected, more consistent effects were found for video plus online games than for either television or (especially) online games alone. Surprisingly, children in the DVD + Web group also showed consistently greater gains than children in the All Materials group, perhaps because of the demands of All Materials teachers' having to make time for *Cyberchase* activities every day.

Effects on problem solving often appeared to be driven more by the TV series than by the online games, probably because television is designed to serve as the central component of *Cyberchase*, the television series presents models of successful problem solving in the context of compelling narratives, and the television series provides greater explanation of mathematical concepts than the games. Nevertheless, the online games provided a context for children to engage in rich mathematical reasoning that resembled the same sorts of progression that have been documented in formal classroom mathematics.

Data on children's performance while playing online games revealed evidence of transfer of learning, not only from the treatment to our posttest measures, but also from children's experience with one *Cyberchase* medium to another. This points to a significant strength of learning from multiple media: The lessons learned from one medium can be applied to enrich children's experience while learning from a second medium.

Learning from Cyberchase

Past research has shown that, even in the absence of multiple media, sustained viewing of the *Cyberchase* television series can produce a significant impact on both the process of children's mathematical problem solving and the sophistication of their solutions. As one might expect, these effects were found most consistently in tasks that were taken directly from television episodes that the children had seen, followed by near transfer and far transfer tasks, respectively

(Fisch, 2003). With that in mind, the present study focused on far transfer, where ceiling effects were least likely to occur. It was designed to replicate the prior findings regarding far transfer from television, and extend them by determining whether transfer effects might arise more strongly through exposure to multiple media.

To that end, data from pretest-posttest problem-solving tasks (both paper-and-pencil and hands-on) were analyzed via a series of model-fitting analyses. Primarily, these consisted of general linear modeling, or GLM.[3] The main questions addressed in each analysis were whether children who used *Cyberchase* media demonstrated greater pretest-posttest gains than children in the No Exposure group, and whether the gains observed among *Cyberchase* children differed as a function of which combinations of *Cyberchase* media they had used.

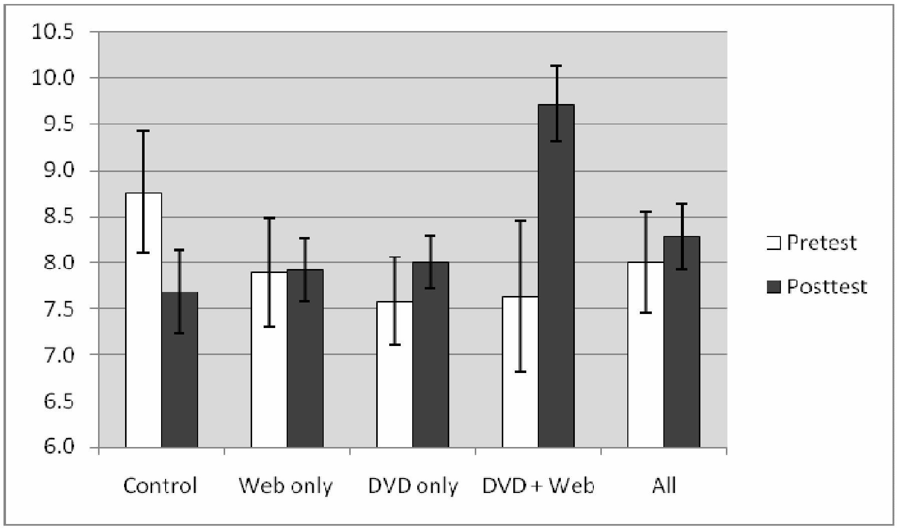
Paper-and-pencil tasks: As noted earlier, children were given two sets of pretest-posttest tasks, one concerning measurement (comparing several roads to find the longest or shortest route) and the other involving organizing data (predicting from data). To ensure that any observed differences among the experimental groups truly reflected effects of the treatment, our analysis controlled statistically for children's reported use of the *Cyberchase* television series and Web site prior to the study.[4]

The following figure presents pretest-posttest change in the sophistication of children's solutions to the organizing data task. For details on what different solution scores represent, please see the sample coding scheme in Appendix C.

[3] A statistical note: Because intact classrooms of children (rather than individual children) were assigned to experimental groups, it was possible that apparent differences among experimental groups might actually be attributable to natural variation between the children's classes. That is, the results could have been influenced by the effects of *nested data*. To rule out such effects, we planned to analyze the data via hierarchical linear modeling (HLM) that would control for nested data. However, preliminary analysis revealed that minimal variability existed between classrooms – too little for HLM analyses to be either possible or necessary. Thus, the data from the experimental phase were not nested data, and could be analyzed via GLM and other methods instead.

[4] This level of control was possible for the paper-and-pencil tasks, because children completed paper-and-pencil tasks individually. However, we could not implement the same control for the hands-on tasks because children worked on hands-on tasks in groups of three, and the triads were not homogenous in terms of their use of *Cyberchase* prior to the study.

Organizing data task (paper-and-pencil):
Mean solution score $\pm$ 1 standard error

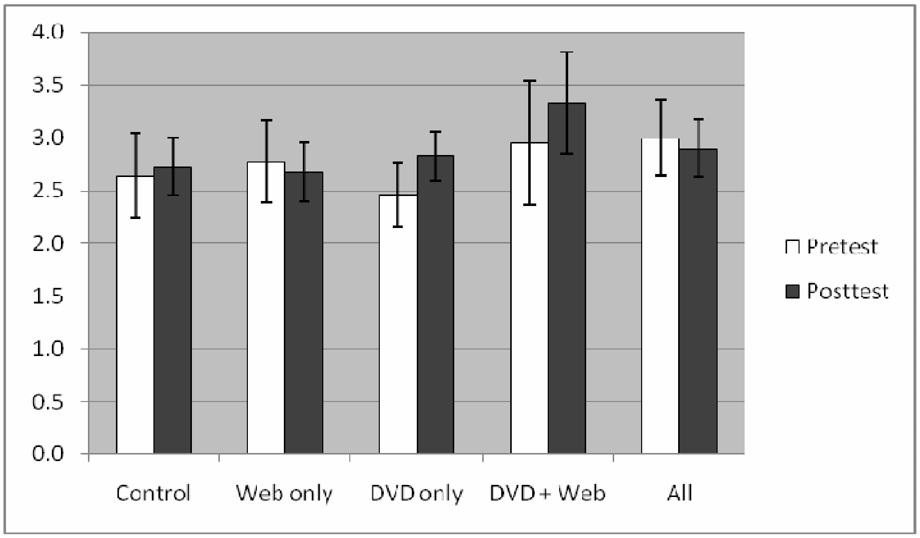


As this figure illustrates, there was a significant effect of *Cyberchase* ($F_{4,295} = 12.13$, $p < .0001$), in that all four *Cyberchase* groups produced more sophisticated solutions in the posttest, while the No Exposure declined from pretest to posttest ($t_{295} = 4.68$, $p < .0001$ for DVD Only vs. No Exposure; $t_{295} = 3.36$, $p < .001$ for Web Only vs. No Exposure; $t_{295} = 6.55$, $p < .0001$ for DVD + Web vs. No Exposure; $t_{295} = 2.18$, $p < .05$ for All Materials vs. No Exposure). For this reason, although few significant differences existed between the groups in the pretest, all of the *Cyberchase* groups scored significantly higher than the No Exposure group in the posttest ($t_{324} = 2.94$, $p < .005$ for DVD Only vs. No Exposure; $t_{324} = 2.35$, $p < .05$ for Web Only vs. No Exposure; $t_{324} = 7.92$, $p < .0001$ for DVD + Web vs. No Exposure; $t_{324} = 2.07$, $p < .05$ for All Materials vs. No Exposure). The effects of *Cyberchase* held constant across boys and girls of different levels of mathematics ability, and across children who either chose or did not choose math as their favorite subject.

With regard to comparisons among children who used different combinations of *Cyberchase* media, as expected, the DVD + Web group improved significantly more than both the DVD Only ($t_{295} = 2.93$, $p < .005$) and Web Only groups ($t_{295} = 3.63$, $p < .0005$). Surprisingly, however, the DVD Only and DVD + Web groups also improved significantly more than the All Materials group ($t_{295} = 2.31$, $p < .05$ and $t_{295} = 4.15$, $p < .0001$, respectively), even though the All Materials group had used more of the *Cyberchase* media.

In the measurement task, too, significant differences emerged among the treatment groups ($F_{4,293} = 2.68$, $p < .05$), but the results for this pair of tasks were less clear-cut. The following figure summarizes pretest-posttest change in the sophistication of children's solutions to the measurement task. Again, detail on what solution scores represent can be found in the sample coding scheme in Appendix C.

Measurement task (paper-and-pencil):
Mean solution score $\pm$ 1 standard error



In this case, the DVD Only and DVD + Web groups improved from pretest to posttest, while the Web Only group remained constant and the All Materials and No Exposure groups declined. (In fact, the All Materials group declined marginally more than the No Exposure group; $t_{293} = 1.80$, $p < .10$.) As a result, the DVD Only and DVD + Web groups both scored significantly higher than the No Exposure group in the posttest ($t_{321} = 2.06$, $p < .05$ and $t_{321} = 2.39$, $p < .05$, respectively). These effects held constant across boys and girls of different levels of mathematics ability, and across children who either chose or did not choose math as their favorite subject.

As in the organizing data task, an unexpected finding arose in that the DVD Only and DVD + Web groups both improved significantly more than the All Materials group ($t_{293} = 3.15$, $p < .005$ and $t_{293} = 2.18$, $p < .05$, respectively). The DVD Only group also improved marginally more than the Web Only group ($t_{293} = 1.64$, $p = .10$).

It is interesting to note that effects of *Cyberchase* emerged more strongly and consistently for organizing data tasks than for measurement, because past research on the *Cyberchase* television series also found more consistent effects for tasks that involved organizing data (Fisch, 2003). It is not clear whether educational content regarding organizing data was conveyed better than measurement, or whether children's understanding of (and misconceptions about) measurement is simply more resistant to change.

Hands-on tasks – process of problem solving: Like the paper-and-pencil tasks, parallel sets of hands-on tasks were administered in the experimental pretest and posttest. Children worked on two sets of hands-on problem-solving tasks, one concerning measurement ("body math" and proportional reasoning) and the other dealing with organizing data (combinatorics and scheduling). Each task was comprised of two parts; in the first part, children attempted to solve the presented problem, and in the second, they were asked to abstract the process they had used to describe a procedure that could be used to solve other, similar problems in the future (similar to the output of the thought-revealing tasks described by Lesh et al, 2000, 2007).

As in the paper-and-pencil tasks, children's solutions to each task were scored on the basis of their level of mathematical sophistication; data regarding children's solutions will be presented in the "Hands-On Tasks – Sophistication of Solutions" section below. In addition, because researchers observed children's process of mathematical problem-solving as they worked on each hands-on task, children were assigned a process score that reflected the number and variety of strategies and heuristics they used to address the problem (e.g., nonstandard measurement, trial and error).[5]

To illustrate the coding used for process and solutions, and children's approaches to solving the problems, the table on pages 36-38 presents an example of one group of three children working on the posttest body math task. In this task, children pretend to be sculptors at a wax museum, who are asked to figure out the dimensions of Shaquille O'Neal's body, based on a photo and outline of his footprint.

Model-fitting analyses were used to compare the experimental groups' growth in each of these types of scores from pretest to posttest. Following the example in the table, we will present the results of these quantitative statistical analyses.

---

[5] Note that, because children worked on these tasks in groups of three, the unit of analysis for hands-on tasks was the triad rather than the individual child. For each task, one process score and one solution score were assigned to the group of three children. Thus, the effective sample size for analyses of hands-on tasks was approximately one-third of the size of the sample for the paper-and-pencil tasks (which each child completed individually).

Sample Group Problem Solving (Shaquille O'Neal task)
Three students: 2 boys, one girl – "Max", "Sam" and "Jeannie" (pseudonyms)

Observations	Coding
Part 1:	
Task is read to the students *Shaq sent his print and picture, so we need to use those to be able to make Shaq's statue the right size, all body parts need to match*.	
Jeannie: "Wow, his foot is bigger than his picture"	Making a comparative observation about size of foot. *(Estimate, approximate)*
Sam measures the foot with a piece of string Jeannie and Max measure with rulers. They discuss whether to measure in cm or inches. They use inches and centimeter interchangeably when discussing their measures. They get 12 inches long and 4 ½ inches across. When discussing their measures they then check the ruler again to settle whether the measurement is in cm or inches.	Students measure with different tools, with string (nonstandard), and with ruler (standard). They use standard units of measure (cm and inches) when measuring. *(Measure: non-standard; Measure: standard)* They measure same object multiple times, checking and rechecking each others' measures. *(Measure: ruler)*
Sam: "Let's times the length of this foot with the length of the picture" Jeannie: "What!? How is that going to do any good? Sam: "I don't know"	Attempting to establish a math connection between Shaq's foot and picture *(Trial and error/guess & check)*
Jeannie: "Okay" (grabs piece of paper and writes 6 ½ x 12) I don't know how to time 6 and a half times twelve.	Questioning reasonableness *(Reasonableness)*
Sam: (grabs calculator and punches in the numbers) announces it's 72 and a half Max: "So he's 7 foot tall"	Use paper-pencil and calculator to compute *(Calculate)*
Jeannie: "Can we cut out the foot?" Max: (grabs scissors and cuts the outline of Shaq's foot)	Manipulate and transform objects *(Manipulate: change object)*
While Max cuts the foot, Jeannie and Sam measure themselves with the ruler. They iterate the ruler and determine their height by 'folding' the ruler (flipping it) upwards. Jeannie measures Sam and announces he is	Use ruler to measure their own heights *(Measure: ruler)*

4 1/2 rulers tall. Measures herself the same way and determines she is 4 rulers tall. Max: "Why are we measuring ourselves?" Sam: I don't know Max: Man we're wasting paper! Killing trees! You should make Shaq come here!	Reasonableness (*Reasonableness*)
Sam: I wonder how big my foot is compared to his? Everyone places their foot on the print to compare Jeannie: I'm probably by his waist	Comparative measuring (*Estimate*)
When asked to share what they have been able to figure out so far: Students say that the foot length is 12 inches long, and that in order to calculate the height they multiplied 12 by either 6 or 7, so the height is about 7 feet tall. They attempt to show this height by saying that it's about Sam's height and Jeannie's height put together. So basically length of the foot times 6 (which they do say is 72, they use the 7 to claim that the height is 7 feet tall). The say the reason to multiply by 6 or 7 is because that's the height in the picture. When pressed why they were multiplying by the height of the picture. They were not really able to explain this. Sam says I don't know because we've seen that been done before.	Accurate measurement of foot length Calculate actual height (erroneously) by multiplying length of actual foot by measure of Shaq's height in the picture.
Part 2: *Need your help in doing what you just did to figure out how to work from a footprint and a photograph to make a statue – for anybody, not just Shaq's*	Recalling information (*Recalling information*)
Students do a bit more talking about what they did and organize their ideas. Max keeps questioning the wisdom of Shaq not sending his actual height measurements. Students wrote and shared the following steps to follow in order for the sculptor to	

make anyone's statue when they send in their foot print and their picture. 1. Measure the width and the length of the shoe print and record it. That helps you make their foot. 2. Measure the picture's height. 3. Times it (picture's height) and the length of the shoe 4. Divide the answer by 7 As an aside, Jeannie says you could also measure the picture's arms and legs.	Reasonableness *(Reasonableness)* Writing in note pad *(Write: list, table, chart)* Inaccurate 'algorithm' for calculating a person's actual height given their foot print and a picture of themselves.

Solution score:

Part 1: 5 points - Inference about height based on incorrect body math

Award 3 points for at least one instance of using standard or nonstandard measurement accurately to document size of footprint, handprint, or hat (within ½" margin of error). And, 2 points for at least one inaccurate use of body math to infer size of other body parts or overall height (without any accurate uses of body math at other points during the task).

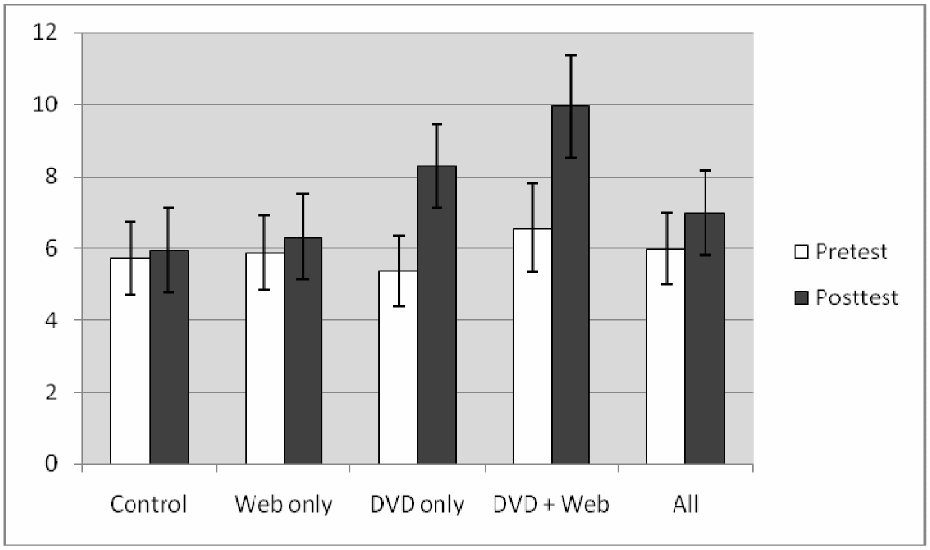
Part 2: 4 points – Inaccurate attempt to use body math

Award 4 points for at least one inaccurate or incomplete attempt at general principle that uses body math to draw inferences about other body parts (e.g., "Measure the foot and multiply by 10 to figure out how tall the person is," "Measure your size [relations] and the other person's, and if it's the same for you, it'll probably be the same for the statue"), without any accurate attempts.

Model-fitting analyses regarding both sets of hands-on tasks revealed that learning from *Cyberchase* was not simply manifest in children's doing a greater number of things while working on the tasks ($F_{4,15}$ = 2.01, N.S. for organizing data tasks, and $F_{4,21}$ = 0.8, N.S. for body math tasks). Rather, the effects of *Cyberchase* appeared in children's using a greater *variety* of strategies and heuristics from pretest to posttest, and in using those strategies and heuristics more effectively.

The following figure presents pretest-posttest change in children's process scores for the hands-on organizing data tasks. The scores in this figure represent the variety of strategies and heuristics children used, via the number of unique heuristics they employed (i.e. not counting duplication if a triad used the same type of heuristic more than once).

Organizing data task (hands-on):
Mean number of unique heuristics used ± 1 standard error



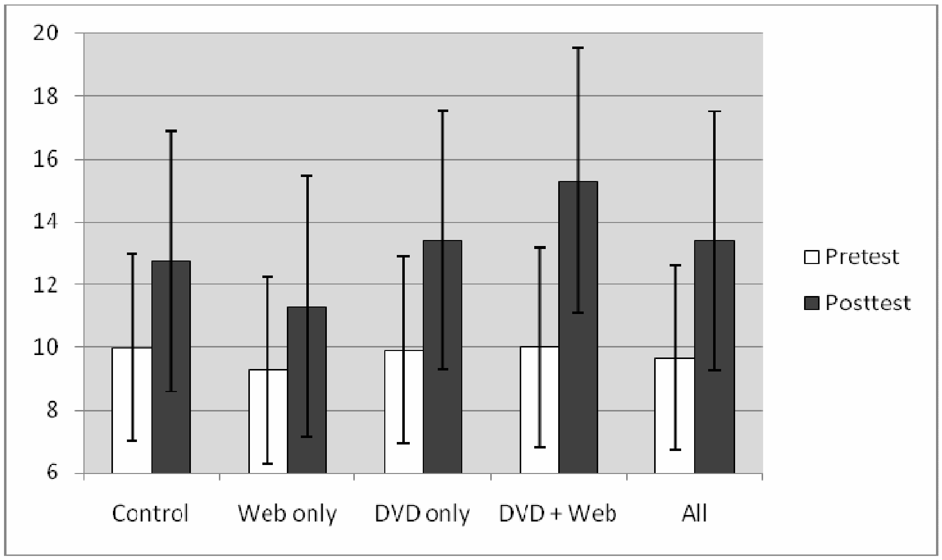
As this figure suggests, a significant effect of *Cyberchase* emerged ($F_{4,19} = 4.21$, $p = .01$), in that the DVD Only and DVD + Web groups both increased significantly more than the No Exposure group in the variety of problem-solving heuristics they applied to the task ($t_{19} = 3.31$, $p < .005$ for DVD Only vs. No Exposure; $t_{18} = 2.97$, $p < .01$ for DVD + Web vs. No Exposure). In the posttest, the DVD + Web group used a significantly greater variety of heuristics than the No Exposure group ($t_{17} = 2.10$, $p < .05$).

For this reason, although few significant differences existed between groups in the pretest, all of the *Cyberchase* groups scored significantly higher than the No Exposure group in the posttest ($t_{324} = 2.97$, $p < .005$ for DVD Only vs. No Exposure; $t_{324} = 2.19$, $p < .05$ for Web Only vs. No Exposure; $t_{324} = 7.88$, $p < .0001$ for DVD + Web vs. No Exposure; $t_{324} = 2.13$, $p < .05$ for All Materials vs. No Exposure). The effects of *Cyberchase* held constant across boys and girls of different levels of mathematics ability, and across children who either chose or did not choose math as their favorite subject.

Turning to comparisons among children who used various combinations of *Cyberchase* media, the DVD Only and DVD + Web groups produced significantly greater gains than either the Web Only ($t_{20} = 2.34$, $p < .05$ for DVD Only vs. Web Only; $t_{18} = 2.32$, $p < .05$ for DVD + Web vs. Web Only) or All Materials groups ($t_{17} = 2.13$, $p < .05$ for DVD Only vs. All Materials; $t_{17} = 2.12$, $p < .05$ for DVD + Web vs. All Materials). In the posttest, the DVD Only, DVD + Web, and All Materials groups all used a greater variety of heuristics than the Web Only group did ($t_{18} = 2.30$, $p < .05$ for DVD Only; $t_{17} = 3.05$, $p < .01$ for DVD + Web, and $t_{19} = 2.25$, $p < .05$ for All Materials).

As in the paper-and-pencil tasks, differences between users and non-users of *Cyberchase* emerged more strongly in organizing data tasks than in measurement tasks. The following figure presents pretest-posttest change in process scores for the body math tasks (again reflecting the unique heuristics that children used, without counting duplication).

Measurement/body math task (hands-on):
Mean number of unique heuristics used ± 1 standard error



As this figure suggests, the overall effect of treatment group on pretest-posttest change was not strong enough to reach statistical significance ($F_{4,22}$ = 1.37, N.S.). However, within groups, the DVD + Web group was the only group that used a significantly greater variety of heuristics in the posttest than they had used in the pretest (t_4 = 3.35, p < .05). As a result, in the posttest, the DVD + Web group used a significantly greater variety than the No Exposure group (t_{17} = 2.10, p = .05).

In addition, the DVD + Web, DVD Only, and All Materials groups all used a significantly greater variety than the Web Only group in the posttest, although no such differences appeared in the pretest (t_{17} = 3.05, p < .01; t_{18} = 2.30, p < .05; and t_{19} = 2.25, p < .05, respectively)

Together, then, the organizing data tasks suggest that children used a greater variety of strategies and heuristics as a result of their use of *Cyberchase*. The measurement tasks provide some additional support for this conclusion in the case of the DVD + Web group. Given this trend, it would be natural to ask which heuristics were primarily responsible for this greater variety. To find out, separate model-fitting analyses were conducted for each of the heuristics in the coding scheme. These analyses revealed significant or marginally significant effects regarding eight heuristics. Most of these effects did not form a coherent trend across heuristics. However, two of them were meaningful:

- *Reasonableness (i.e. reconsiders own ideas):* A marginally significant overall effect was found for this heuristic in the measurement task ($F_{4,341}$ = 2.01, p < .10). All four

Cyberchase groups were either significantly or marginally more likely to evaluate their own ideas in the posttest, whereas the No Exposure group showed no change.

- *Recalls Cyberchase (i.e. spontaneously recalls relevant information from Cyberchase and labels it as such):* A marginally significant effect was found for this heuristic in the measurement task ($F_{4,339} = 2.12$, $p < .10$). In the posttest, all four *Cyberchase* groups explicitly (and spontaneously) recalled information from *Cyberchase* to help them with the task, whereas the No Exposure group did not.

Thus, in the posttest, children who used *Cyberchase* were somewhat more likely to question and reconsider their own ideas as they worked on the measurement task. This is consistent with our qualitative observations that, in the posttest, *Cyberchase* users employed top-down approaches to problem solving and demonstrated persistence when one attempt at a solution did not succeed (as will be discussed later, in the section on "Qualitative Observations – Hands-On Problem Solving"). It is not surprising that *Cyberchase* users explicitly recalled information from *Cyberchase* more often than the No Exposure group (although even the No Exposure group could have done so, if they had ever encountered *Cyberchase* outside the context of the study). However, the fact that approximately one-third of users spontaneously referred to *Cyberchase* while working on posttest tasks supports the conclusion that the observed pretest-posttest differences were indeed attributable to their experience with *Cyberchase*.

As noted, the remaining significant or marginally significant differences in use of individual heuristics did not form a cohesive trend across the heuristics:

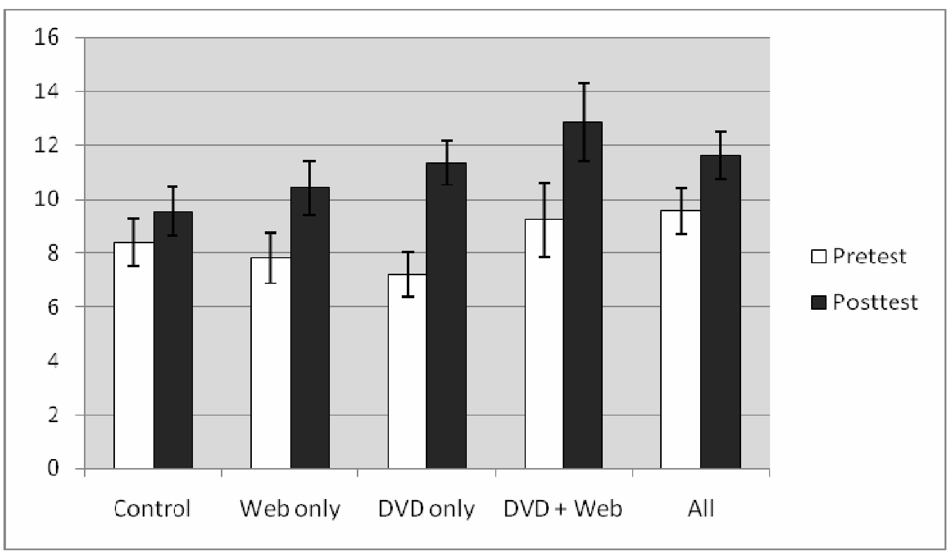
- *Manipulate: Change objects (e.g., build a nonstandard measuring device, make notations directly on hands-on materials):* A significant effect was found for this heuristic in the measurement task ($F_{4,341} = 5.30$, $p < .0005$). The All Materials group increased in their use of the heuristic from pretest to posttest, while the No Exposure and DVD Only groups used it less often in the posttest.
- *Manipulate: Use objects (e.g. use one object to represent another, manipulate cards to physically make combinations):* Significant effects were found regarding this heuristic in both sets of tasks ($F_{4,344} = 2.41$, $p < .05$ for organizing data, and $F_{4,341} = 2.79$, $p < .05$ for measurement). In the posttest organizing data task, the DVD Only group used this heuristic significantly more than the No Exposure group, and the DVD + Web group used it significantly more than the Web Only group. In the measurement task, the All Materials group used this heuristic less from pretest to posttest, while the No Exposure group used it more in the posttest.
- *Gather information (e.g., observing or studying hands-on materials):* Significant effects were found regarding this heuristic in both sets of tasks ($F_{4,341} = 6.62$, $p < .0001$ for organizing data, and $F_{4,344} = 2.01$, $p < .10$ for measurement). In both cases, the DVD + Web group used this heuristic either significantly or marginally more in the posttest than the pretest. The No Exposure showed a similar gain in the measurement task, but did not change in the organizing data task. The All Materials group spent less time gathering information in the measurement task posttest, and the Web Only group showed a marginally significant decrease in the organizing data task.

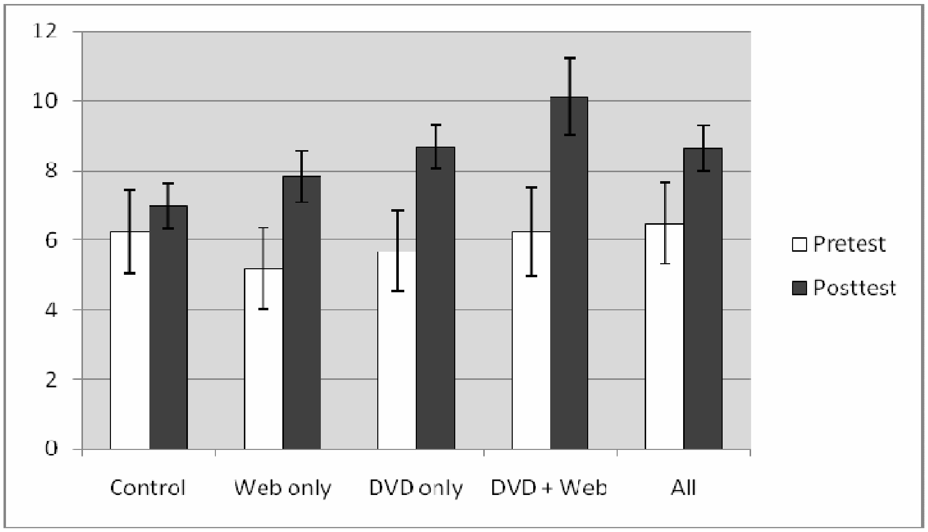
- *Write: List (i.e. record data in a list to keep track of it):* A significant effect was found for this heuristic in the organizing data task ($F_{4,344} = 3.84$, $p < .005$), in that the All Materials group wrote lists more often from pretest to posttest, while the DVD Only group showed a small decrease.
- *Estimate:* A marginally significant effect was found for this heuristic in the measurement task ($F_{4,341} = 2.00$, $p < .10$). The Web Only group used significantly more estimation in the posttest than in the pretest.
- *Look for patterns (e.g., in proportional relationships among body parts):* A significant effect was found for this heuristic in the measurement task ($F_{4,341} = 2.20$, $p < .0005$), in that the Web Only group used it slightly less in the posttest, whereas all of the other groups increased.

Hands-on tasks – sophistication of solutions: As noted earlier, children's solutions to each hands-on task were comprised of two parts; the first part asked them to solve the presented problem, and the second part asked them to abstract their work to describe a way in which other, similar problems could be addressed in the future. The sophistication of the children's answers to each part was scored separately, and also pooled into a total solution score that reflected the sophistication of their solution to the task as a whole.

The following figures show pretest-posttest change in children's solutions to the two sets of tasks.

Organizing data task (hands-on):
Mean solution score ± 1 standard error





Measurement/body math task (hands-on):
Mean solution score $\pm$ 1 standard error

In both sets of tasks, all five treatment groups showed improvement from pretest to posttest. As a result, neither set of tasks elicited significant overall differences among the treatment groups ($F_{4,20} = 1.27$, *n.s.* for organizing data; $F_{4,19} = 1.71$, *n.s.* for measurement). However, pairwise comparisons in the organizing data tasks revealed that the DVD Only group improved significantly more than the No Exposure group from pretest to posttest ($t_{21} = 2.17$ $p < .05$).

In addition, finer-grained analyses revealed several other significant effects that favored some or all of the groups that used *Cyberchase*. In the organizing data tasks, within-group analysis found that the DVD Only and Web Only groups both showed significant gains from pretest to posttest ($t_6 = 4.21$, $p < .01$, and $t_{11} = 2.20$, $p = .05$, respectively). The same was true for the All Materials group, although their gains were only marginally significant ($t_5 = 2.18$, $p < .10$). Neither the No Exposure nor (surprisingly) the DVD + Web group showed significant gains ($t_6 = 0.96$, *n.s.*, and $t_{12} = 1.61$, *n.s.*, respectively). In the posttest, the DVD + Web and All Materials groups both produced marginally more sophisticated solutions than the No Exposure group ($t_{21} = 1.97$, $p < .10$, and $t_{23} = 1.68$, $p < .10$, respectively).

In the measurement tasks, the DVD + Web group was the only one whose within-group gains were large enough to be statistically significant ($t_6 = 2.85$, $p < .05$). As a result, in the posttest, the DVD + Web group produced significantly more sophisticated solutions than the No Exposure group ($t_{20} = 2.49$, $p < .05$). The posttest solutions produced by the DVD Only and All Materials groups were also somewhat more sophisticated than the No Exposure group's, although this difference was only marginally significant ($t_{21} = 1.89$, $p < .10$, and $t_{22} = 1.82$, $p < .10$, respectively).

In both sets of tasks, effects on solution were centered primarily in the second part of the task, in which children had to abstract their work to describe a way in which other, similar problems could be addressed in the future. In the organizing data tasks, within-group analysis found that three out of the four *Cyberchase* groups improved significantly on this part of the task (t_{12} =

3.95, $p < .005$ for DVD Only; $t_{25} = 2.52$, $p < .05$ for Web Only, and $t_{14} = 3.43$, $p < .005$ for All Materials), whereas the No Exposure and DVD + Web groups did not ($t_{26} = 0.74$, *n.s.* for No Exposure, and $t_{17} = 1.50$, *n.s.* for DVD + Web). Similarly, in the measurement tasks, all four *Cyberchase* groups improved significantly from pretest to posttest ($t_{20} = 3.76$, $p = .001$ for DVD Only; $t_{27} = 2.65$, $p = .01$ for Web Only; $t_{19} = 3.26$, $p < .005$ for DVD + Web; and $t_{20} = 3.03$, $p < .01$ for All Materials), whereas the No Exposure group did not ($t_{26} = 0.99$, *n.s.*). The emergence of significant effects in this part of the tasks is especially interesting, both because this part of the tasks was closest to the design of the thought-revealing activities discussed by Lesh et al. (2000, 2007) and others, and because its focus on describing a process to be applied to future problems is particularly relevant to transfer of learning.

If we consider the data regarding process and solution scores (both hands-on and paper-and-pencil) together, three noteworthy trends are evident. First, there were numerous instances in which one or more of the groups that used *Cyberchase* outperformed the No Exposure group, speaking to the educational power of *Cyberchase* as a whole. Second, as expected, effects often appeared more consistently in the DVD + Web group than in either the DVD Only or Web Only groups, suggesting greater learning from multiple media than from either medium alone. Surprisingly, though, children in the DVD + Web group also showed consistently greater gains than children in the All Materials group (which used the same materials plus hands-on classroom activities). Although we cannot be certain, we believe that the less consistent performance of the All Materials group may have been influenced by attitudinal cues in teachers' behavior in response to the demands of having to make time for *Cyberchase* activities every day.

Third, effects on problem solving often appeared to be driven more by the television series than by the online games, as seen in more consistent effects in the DVD Only group (and other groups that viewed the television series) than in the Web Only group. We suspect that this is due to the fact that television is designed to serve as the central component of *Cyberchase*, embeds mathematical content and successful problem solving in the context of compelling stories, and provides greater explanation of mathematical concepts than the games (which allow children opportunities to exercise skills, but present less overt explanation). We will return to all three of these points in the Conclusions and Implications section at the end of this report.

Problem solving during online games: Although the *Cyberchase* television series produced stronger pretest-posttest effects than the online games, this is not to say that the online games were without educational value. Elsewhere, we have drawn on pilot data gathered in support of the present study to examine children's mathematical problem-solving while playing several online *Cyberchase* games (Fisch et al., in press). Analyses of both in-person observations and online tracking data of children's clicks pointed to the games' providing a context for rich mathematical reasoning. While playing *Cyberchase* online games, children engaged in processes of problem solving that resembled the sorts of processes that past research has documented during mathematical problem solving in formal education.

Just as one might expect in offline mathematical reasoning, we found a range of sophistication in the mathematical strategies that children used while playing the games. Moreover, parallel to research on classroom mathematics (e.g., Lesh et al., 2000) and findings within the developmental literature on children's strategy usage (e.g., Siegler, 2007), those children who used more sophisticated strategies often did not apply them immediately. Rather, they engaged in cycles of problem solving that began with less sophisticated strategies and progressed to more sophisticated approaches when necessary.

This finding was confirmed in the present, full study. For example, in the full study, 145 children played the game Railroad Repair at least once. While playing the game for the first time, 68% of these children showed evidence of at least one use of the highest-level strategy in our coding scheme, 28% progressed as far as a mid-level strategy, 1% never moved beyond the most basic strategy, and 2% did not employ any of these strategies (and, as a result, did not provide any correct answers). Yet, despite the fact that 96% of the children used more sophisticated strategies at some point during the game, all but two of the 145 children used the most basic strategy at the beginning of the game. Thus, parallel to classroom reasoning, even those children who were capable of more sophisticated reasoning typically began by using a more basic strategy (perhaps in the interest of efficiency), and the level of their reasoning evolved over the course of gameplay, in response to the increasing demands of the game.

Qualitative observations – hands-on problem solving: Whereas quantitative, statistical analysis has great value in providing the methodological rigor necessary to evaluate children's learning, it also can be somewhat limited in its ability to capture the flavor of such learning. To better understand the nature of the learning that produced the significant effects documented above, we now turn to qualitative observations of children's hands-on problem solving. The goal of this section is to describe the nature of the pretest-posttest changes observed in children's problem solving, and to suggest several plausible conjectures about possible causes of these changes.

To appreciate the significance of the changes that were observed, it is important to go beyond wishful thinking and recognize the current state of research on problem solving in mathematics education. Every ten years, the *National Council of Teachers of Mathematics* publishes a *Handbook of Research in Mathematics Education* (Lester, 2007); the 2007 edition was edited by Frank Lester, one of the foremost, nationally prominent researchers focusing on mathematical problem solving. As plans were being formulated for this huge project, Lester noted that research on problem solving has declined significantly during the past decade. So, he specifically asked the authors of the chapter on problem solving (i.e., Richard Lesh and Judi Zawojewski) to analyze these trends, to explain the most significant causes and trends, and to describe promising new directions for future research. The following facts were noted.

- Polya-style problem solving heuristics – such as *draw a picture*, *work backwards*, *look for a similar problem*, or *identify the givens and goals* - have long histories of being advocated as important abilities for students to develop (e.g., Polya, 1957). But, it is not at all clear what it mean to "understand" them. Such strategies clearly have descriptive power. That is, experts often use such terms when they give after-the-fact explanations of their own problem solving behaviors - or those of other people that they observe. Nonetheless, the state of research on mathematical problem solving has not changed significantly since Begel's comprehensive review of the literature in 1979:

    "There is little evidence that general processes that experts use to describe their past problem solving behaviors should also serve well as prescriptions to guide novices' next-

> steps during ongoing problem solving sessions…[N]o clear cut directions for mathematics education are provided by the findings of these studies. In fact, there are enough indications that problem-solving strategies are both problem- and student-specific often enough to suggest that hopes of finding one [or a few strategies] which should be taught to all [or most students] are far too simplistic.” (Begel, 1979, p. 145)

- Similarly, Schoenfeld’s (1992) review of the literature again concluded that attempts to teach students to use general problem-solving strategies have not been successful. Schoenfeld noted that Polya’s descriptive processes are really more like names for large categories of processes rather than being well defined processes in themselves. So, wishful thinking not withstanding, short lists of descriptive processes tend to be too general to be useful; yet, long lists of prescriptive processes tend to become so numerous that knowing when to use them becomes the heart of understanding them.

In the light of the preceding facts about the state of research on mathematical problem, the point that is perhaps most noteworthy from this *Cyberchase* assessment project is that *students did in fact improve significantly in their problem solving capabilities*. Furthermore, an on-site visitor to our research sites would have provided evidence that would be even more impressive than the results that are clear in the project’s statistical analyses.

In the observations that follow, special attention will be given to students’ performances on two pairs of 20-minute, hands-on tasks that were used as pre- and posttest items. As discussed in the Method section earlier in this report, the pretest problems were called the *Ping Pong* and the *Detective tasks*; and the posttest problems were called the *Soccer* and the *Big Foot tasks* (or *Shaquille O’Neal task*). All four problems are junior versions of *model-eliciting activities* that have been used in a number of research projects investigating models and modeling abilities for students from elementary school through high school, college, and graduate students – as well as in-service and pre-service teachers, and professionals in fields such as engineering were mathematical thinking is an important component of expertise (Lesh & Doerr, 2002; Lesh, Hamilton & Kaput, 2007). Such activities were designed explicitly to be simulations of “real life” problem solving situations – and to focus on elementary-but-deep mathematical concepts and abilities that are seldom addressed on short-answer standardized tests. Consequently, such activities are especially useful for identifying high-ability students who seldom emerge as being highly capable using traditional tests.

Of course, doing “bean counts” of process objectives is problematic for a variety of reasons. For instance, it often was difficult for coders to determine whether a single process was used several times, or whether these should be counted as different instances of process use. Similarly, simple counts do not take into account the fact the different processes should perhaps be given different weights. Nonetheless, the fact that (as noted earlier) *Cyberchase* users demonstrated significantly greater growth than non-users in both the variety of heuristics they used and the sophistication of their solutions suggests that what changed from early activities to later activities was the size, or power, or effectiveness of processes used – and not gross number of processes-strategies-ideas used.

What appeared to explain these differences in effectiveness with which processes, strategies, and ideas were used? For the researchers and research assistants who were on-site for the start-of-project and end-of-project problem solving sessions, a number of changes were obvious in children's behaviors. In Indiana, some of these behaviors were especially striking because most of the research team was also working on a separate research project in which undergraduate students and graduate students were observed working on model-eliciting activities that were very similar to those that were used with children in the *Cyberchase* project. So, for these research assistants, the following facts were obvious in posttest problem solving performance for children who had participated in the project:

- *The children were remarkably effective at working well in groups.* For example, even compared with more mature students at the university, a variety of diverse and appropriate roles were adopted by children in the groups; few students were left out; and roles often shifted appropriately when needs arose. In doing so, the children often referred to *Cyberchase* stories that they had watched – and to roles played by the *Cyberchase* characters.

- *The children often actually engaged in "top down" planning!* By contrast, no instances of such behavior were recorded for problem solving sessions at the start of the project. That is, for posttest problem-solving activities, before charging ahead on a plan (with little reflection), the children often did something akin to "brainstorming" and formulated a general plan before they proceeded along some presumed solution path. *This is very uncommon behavior* – even among students we have observed at the university. And, it seemed to be the most important single factor that explained performance increases for students who viewed the *Cyberchase* television series. Again, references to *Cyberchase* stories seemed to provide the most obvious stimulus and guide for these behaviors.

- *The children persevered.* They were not disturbed if their first way of thinking about the problem didn't work, and they assumed from the start that the problem was not going to be solved in a couple of minutes. Again, references to *Cyberchase* episodes seemed to provide models of the kinds of behavior that would be needed.

Whereas programs that aim at increasing problem solving performance traditionally focus on teaching processes, strategies, heuristics, beliefs, dispositions, or attitudes that are embodied in rules – in the hopes that students will later connect these rules to relevant concepts needed to solve problems – the primary factor that seemed to explain success on this project was "stories." If one takes seriously the research of researchers like Lakoff, Schank, or Lesh, then this should not be surprising. In virtually every field where researchers have investigated differences between effective problem solvers and those who are less effective, it has become clear that effective problem solvers not only *do* things differently, but they also *see* (or interpret) things differently. Relevant theories have described these interpretation systems have been described using the language of models, metaphors, stories, or scripts. But, in any case, they cannot be summarized as fact, a process, or a rule. And, they are not likely to be assessed by standardized tests.

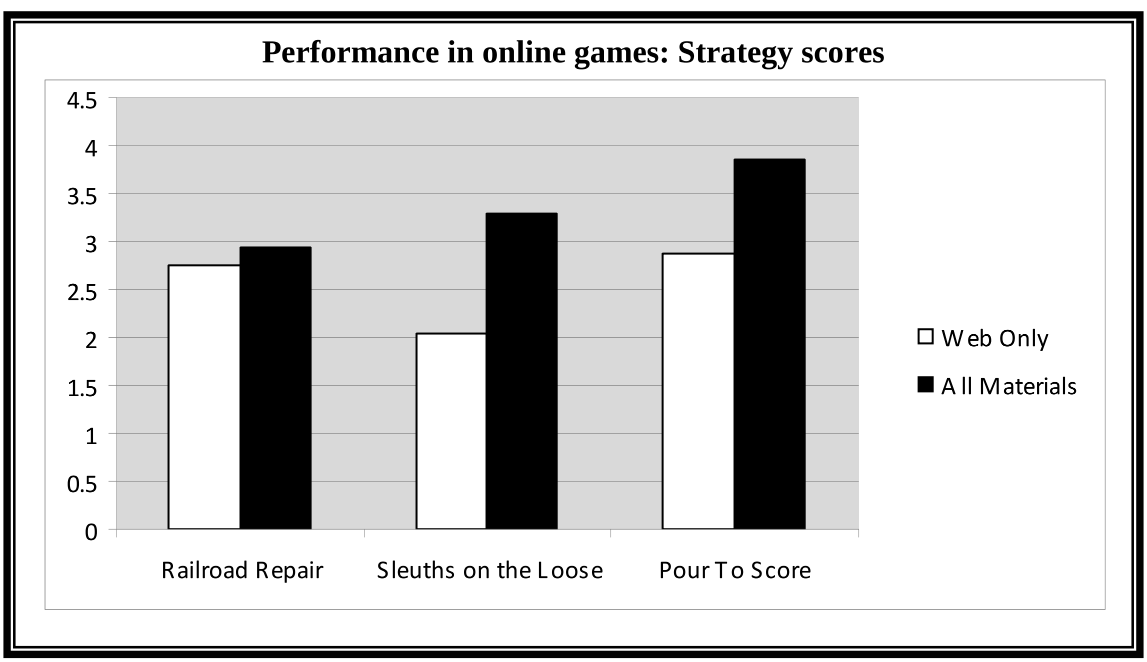
Did these activities allow us to document the abilities of students whose potential had not previously been apparent in the context of traditional word problems of the type emphasize in textbooks and tests? All we can say from the current study is that most of the participating teachers believed that this was the case. That is, we often recorded comments from teachers who said that *"I've never seen [Johnny or Jenny] do this well before in math."* Similarly, children often commented that *"Different characters contributed in different ways [to solutions in Cyberchase episodes]."* (For further comments from teachers and children, see the "Perceptions of *Cyberchase* and Learning" section later in this report.) Again, the adoption of an effective problem-solving persona did appear to be reducible to a pledge of allegiance to some easily stated belief, habit of mind, disposition, or rule of behavior.

Cross-Platform Learning

Earlier, we presented evidence of transfer of learning in terms of children's ability to apply the sorts of problem-solving strategies and heuristics modeled in *Cyberchase* to new problems in the posttest (resulting in stronger pretest-posttest gains among users of *Cyberchase* than among non-users). This sort of transfer is clearly an important part of the impact of any media-based informal education project, whether the project employs several forms of media or just one.

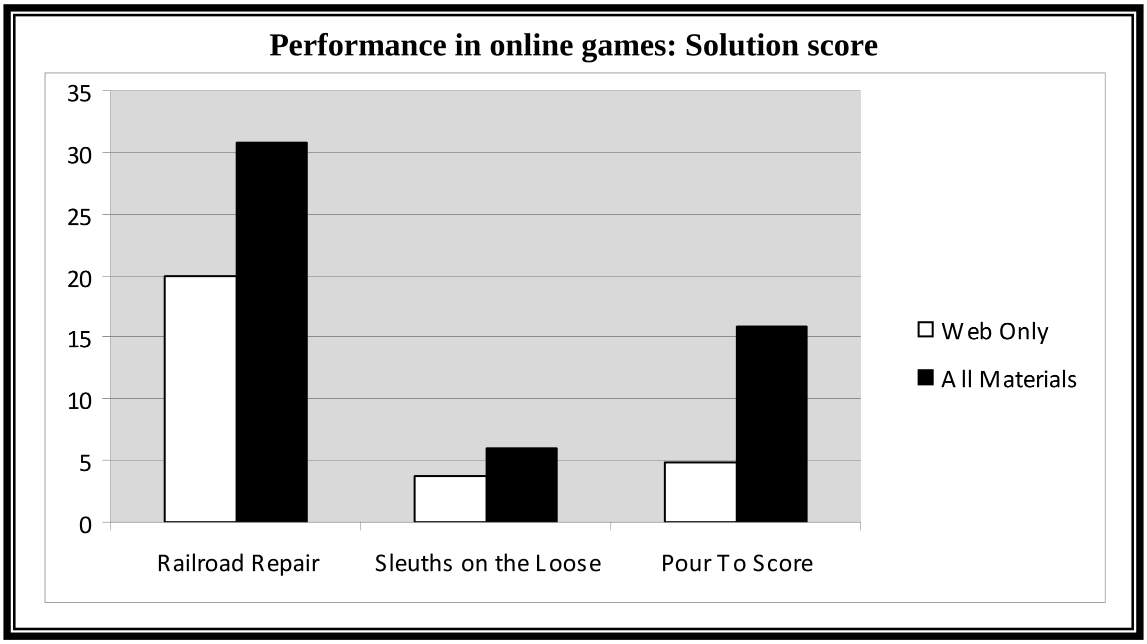
From the standpoint of cross-platform learning, however, we were also interested in how learning from one medium might interact with learning from another. In this respect, we were interested, not only in how learning from one medium might transfer to a posttest assessment, but also in how it might transfer into children's engagement and performance while interacting with (and learning from) a different form of educational media. Specifically, we used the online tracking data discussed above to compare performance in three online *Cyberchase* games between children who only played the games (i.e. the Web Only group) and those who also used other forms of *Cyberchase* media (i.e. the All Materials group). Parallel to the analysis of hands-on tasks discussed earlier, our analysis of online tracking data focused on two aspects of performance in the games: strategy scores (reflecting the sophistication of the strategies children used while playing the game) and solution scores (equivalent to the number of correct responses that children produced in each game). For further information about these scores, see the coding schemes presented in Appendix G.

The following figure compares the strategy scores received by the All Materials and Web Only groups in each of the three games.



On average, children in the All Materials group applied significantly more sophisticated strategies than the Web Only group while playing each of the three games ($t_{109} = 3.04$, $p < .005$ for Railroad Repair; $t_{68} = 3.01$, $p < .005$ for Sleuths on the Loose; and $t_{89} = 4.82$, $p < .0001$ for Pour to Score).

As a result, these children also produced more correct responses in two of the three games, as seen in the following figure.



On average, the All Materials group provided significantly more correct responses than Web Only group while playing Railroad Repair ($t_{93} = 2.72$, $p < .01$) and Pour to Score ($t_{92} = 3.96$, $p < .0001$). The effect on solution scores in Sleuths on the Loose was not strong enough to achieve significance ($t_{68} = 1.282$, *n.s.*).

Taken together, this pattern of online effects points to a significant strength of cross-platform learning: The lessons learned from one medium can be applied, not only to enrich children's general knowledge, but also to enrich children's experience *while they are in the process of learning from a second medium*. These effects are particularly intriguing in the case of Railroad Repair, because its mathematical content (adding decimals) was not aligned closely with any of the television episodes or hands-on games in the treatment, although the same sorts of underlying problem-solving strategies and systematic thinking could be used. We shall return to this point in the Conclusions and Implications section below.

Impact on Attitudes Toward Mathematics and Problem Solving

Overview of Results

Paper-and-pencil measures of attitude revealed only one pair of significant effects: From pretest to posttest, all of the *Cyberchase* groups sustained their interest and (to a lesser degree) confidence in doing school math, while the attitudes of the control group declined. No significant effects appeared for other domains of out-of-school mathematics.

However, we also found behavioral evidence of an effect on children's motivation: In two of the three *Cyberchase* online games, users of multiple media were more likely to continue playing beyond the end of the game than children in the Web Only group, pointing to their greater motivation to engage in a fun, mathematical activity.

Attitude

As described in the Method section, several measures were used to investigate the effects of multiple *Cyberchase* media on children's attitudes toward mathematics and problem solving. One pair of paper-and-pencil attitude scales focused on children's interest and confidence in engaging in a variety of mathematical activities. Motivation was assessed via a paper-and-pencil measure that addressed orientation toward engaging in challenging mathematics activities (mastery vs. performance), and online tracking data provided a behavioral means for gauging motivation toward engaging in math-based games.

In comparison to the problem-solving data presented earlier, fewer significant effects emerged regarding attitude. However, some significant effects did appear, as we shall see.

Interest and confidence: Interest and confidence scales were administered in both the pretest and posttest. Children were asked to rate their interest and confidence in a variety of mathematical and non-mathematical tasks that fell into four categories (and subscales): *Cyberchase* math (i.e. mathematics that was presented in *Cyberchase* materials during the treatment), non-*Cyberchase* math (i.e. out-of-school mathematics that did not appear in any of the *Cyberchase* materials in the treatment), school math (e.g., solving blackboard problems or studying for math tests), and non-math (i.e. non-mathematical activities, such as exploring the history of one's home town). For each activity mentioned in the scale, children rated their interest on a five-point scale (with the most positive response option coded as 5), and rated their confidence on a parallel five-point scale.

The following tables present children's mean responses regarding interest and confidence.

Interest: Mean scores

	Control		DVD Only		Web Only		DVD + Web		All Materials	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
School math	5.47	3.71	3.20	3.46	3.34	3.58	4.18	3.27	3.31	3.33
***Cyberchase* math**	4.13	3.64	3.50	3.61	3.71	3.68	3.96	4.05	3.35	3.48
Non-*Cyberchase* math	4.00	4.26	3.40	3.50	3.45	3.54	---	3.76	3.21	3.43
Non-math	4.42	3.55	3.88	3.57	4.10	3.49	4.30	4.02	3.82	3.47

Confidence: Mean scores

	Control		DVD Only		Web Only		DVD + Web		All Materials	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
School math	4.76	4.59	4.25	4.50	4.31	4.46	4.33	4.83	4.56	4.73
***Cyberchase* math**	4.01	3.74	3.62	3.74	3.61	3.77	3.77	4.09	3.56	3.77
Non-*Cyberchase* math	3.72	4.06	3.59	3.91	3.70	3.95	---	4.30	3.54	3.86
Non-math	4.18	3.99	4.15	3.86	4.22	3.79	4.18	4.15	3.99	3.93

Note: By chance, during the pretest, too few children in the DVD + Web group responded to the non-*Cyberchase* math items to allow for statistical analysis. Thus, this cell has been omitted from each table.

In interpreting these data, it is helpful to consider the results obtained from past research regarding the effects of other mathematics-based television series. Some past studies have found sustained viewing of series such as *Square One TV* to result in increased attitudes (e.g., Debold et al., 1990), but others have found viewing of series such as *Futures* to produce a different sort of effect: helping to sustain children's positive attitudes toward mathematics while nonviewers' attitudes declined (Research Communications Ltd., 1992; cf. Crane et al., 1994).

As the above tables illustrate, the latter type of effect was evident here. In almost every dimension, the No Exposure group's interest and confidence declined from pretest to posttest, and the same was true for non-mathematics activities among all of the treatment groups. By contrast, in all of the mathematics-related dimensions, the four *Cyberchase* groups' ratings remained fairly constant from pretest to posttest. This difference between users and non-users of *Cyberchase*, although evident throughout the data from both scales, achieved statistical significance only for interest in school math ($F_{4,267} = 6.91$, $p < .0001$) and marginal significance for confidence in school math ($F_{4,270} = 2.16$, $p < .10$).

Model-fitting analyses revealed that the No Exposure group declined significantly more than all of the *Cyberchase* groups in interest in school math ($t_{267} = 4.54$, $p < .0001$ for DVD Only; $t_{267} = 4.16$, $p < .0001$ for Web Only; $t_{267} = 2.46$, $p = .01$ for DVD + Web; and $t_{270} = 3.80$, $p < .0005$ for All Materials). Similarly, the No Exposure group also declined either significantly or marginally more than all four *Cyberchase* groups in confidence in school math ($t_{270} = 2.38$, $p < .05$ for DVD Only; $t_{70} = 1.71$, $p < .10$ for Web Only; $t_{270} = 2.81$, $p = .005$ for DVD + Web; and $t_{270} = 1.91$, $p < .10$ for All Materials).

Motivation: No significant effects were found in our paper-and-pencil measure of orientation toward motivation, in any of the three contexts used in the measure: classroom mathematics (χ^2_8 = 3.21, *n.s.*), video games (χ^2_8 = 6.62, *n.s.*), or out-of-school mathematics (χ^2_8 = 2.80, *n.s.*).

However, significant effects did emerge in a behavioral measure of motivation that employed data from the online *Cyberchase* games to assess the number of children who were sufficiently motivated to continue playing a math-based game even after they reached the end of the game. As discussed in the Method section, it is typically difficult to measure motivation via variables such as time on task, because such measures confound the effects of ability and motivation. In the present study, however, we overcame this challenge by restricting our motivation analysis to only those children who reached the end of each online game; since all of the children in the subsample had sufficient ability to successfully complete the game, continuing to play beyond the end of the game was a clear indicator of motivation. Thus, by comparing children who only played the games (the Web Only group) to children who used multiple *Cyberchase* media (the All Materials group), we could determine whether the use of multiple media increased children's motivation to engage further in mathematics-based games.

The following tables present the percentage of children in the All Materials and Web Only groups who chose to either stop or continue playing once they reached the end of a game.

Railroad Repair game

	Continue beyond end	Do not continue
All Materials	32%	68%
Web only	8%	92%

Pour to Score game

	Continue beyond end	Do not continue
All Materials	31%	69%
Web only	0%	100%

Sleuths on the Loose game

	Continue beyond end	Do not continue
All Materials	12%	88%
Web only	15%	85%

Chi-square analysis revealed that, in two of the three games, children in the All Materials group were more likely to continue playing beyond the end of the game than children in the Web Only group (χ^2_1 = 8.09, $p < .005$ for Railroad Repair; χ^2_1 = 7.90, $p = .005$ for Pour to Score; and χ^2_1 = 0.10, *n.s.* for Sleuths on the Loose). Thus, it appeared that children who used multiple *Cyberchase* media were more motivated to engage in a fun, mathematical activity that featured the same characters and world. We believe that this finding holds important implications for the role of multiple media in contributing to children's interest in academic subjects, especially from the perspective of Hidi and Renninger's (2006) four-phase model of interest development. We will return to this point in the Conclusions and Implications section.

Perceptions of *Cyberchase* and Learning: Supplementary Teacher and Child Interviews

Following the experimental posttest, supplemental interviews were conducted with participating children and teachers, to help us interpret the data and understand the effects found in the study. These interviews focused on *Cyberchase* itself: their experiences with and reactions to the materials, and what (if anything) they believed children learned from *Cyberchase*.

Please note that, because these interviews concerned *Cyberchase* itself, they could not be conducted with children in the No Exposure group, who had not used the materials.

Perceptions of learning: The following table shows the percentage of children in each of the treatment groups who believed that they had or had not learned from *Cyberchase*.

Percentage of children who believed they:	All Materials	DVD + Web	DVD Only	Web Only
Learned from Cyberchase	90%	100%	96%	70%
Did not learn	10%	0%	4%	30%

As this table shows, the vast majority of children in all four *Cyberchase* groups felt that they had learned from *Cyberchase*. However, just as the experimental data found that impact was often smaller in the Web Only group, significantly fewer children in this group believed they had learned (although most of the Web Only children said they learned as well). More than 90% of the children in each of the other groups believed they had learned from *Cyberchase*, as opposed to 70% of the children in the Web Only group ($\chi^2_4 = 18.80$, $p < .001$).

Differences also appeared in children's accounts of *what* they thought they had learned from *Cyberchase*, as shown in the following table.

Number of children who believed they learned:	All Materials	DVD + Web	DVD only	Web only	Total
"Math" in general	17	1	3	1	22
Arithmetic	2	5	13	9	29
Problem solving	10	14	19	6	49
Measurement (general)	18	10	34	1	63
Measurement: body math	2	0	6	0	8
Area/perimeter	5	0	9	0	14
Fractions	3	0	1	4	8
Geometry	2	0	0	2	4
Patterns/tesselation	3	0	1	0	4
Numbers/counting	3	0	1	0	4
Graphs	2	0	2	0	4
Estimation	1	0	1	0	2

As this table indicates, one of the most prevalent responses in all four groups was that children learned about problem solving. Among children's more specific responses, however, children in the All Materials, DVD + Web, and DVD Only groups most often discussed measurement (including several specific mentions of body math), whereas the Web Only group most often mentioned arithmetic.

Examples of the children's responses included the following:

> *"[The television episode about body math] showed me how to guess people's height and all of that. I just, previously, I didn't know anything about that. I was just like, "What?" I was like, "Oh! I never knew that." So it really got me thinking and it was really fun trying to see in the next one seeing how tall Shaq was [in the posttest task]. I was close. I was doing the math-y way. I love math."*
>
> *-- Girl, All Materials*

> *"I learned that just 'cause, like, there's a perimeter or something, even if it's the same length, it could be a little different. Like [if] the shorter the width gets… the longer the height gets, then they're the same. It could be the same but a different shape."*
>
> *-- Boy, All Materials*

> *"Like Railroad Repair [game about adding decimals], it helped me get the pieces I need to put in. And Cyber Olympics [game about organizing and using data], that, you get to pick cards, you get to choose wisely. And Bike Route [taxi geometry game], you have to go on these squares and you can only move certain times."*
>
> *-- Boy, Web Only*

> *"It teaches me how to do math, like fractions, division, multiplication, addition, subtraction."*
>
> *-- Girl, Web Only*

Consistent with the children's perceptions of learning, all of the teachers who were interviewed felt that children had learned from *Cyberchase*. Examples of teacher comments included:

> *"Yes, they definitely benefited from using Cyberchase. In particular, they learned to examine problems for ways to solve them. They also worked with data in many different ways and lots of measuring and measurement terms. When looking at specific content, they learned and benefited from practicing with perimeter and area, graphs, estimating and probability."*

> *"My kids obviously are getting better at problem solving. They don't give up as easily - even my kids who usually have a hard time with math."*

> *"At first, some of my parents complained about me asking the kids to watch TV. But, after they saw what [Cyberchase] was about, they didn't mind."*

> *"I'm sure my kids are going to do better on ISTEP [standardized tests]. You can just see*

it. And, they're much better at explaining what they do. Before, they sometimes could do it, but they couldn't explain…We worked at lot on that - explaining what you're doing to others."

"I think kids remember things longer when they associate things with stories. They sometimes say, 'This is like when [the Cyberchase characters did something]."

"There is a child in my room that I know learned from the Web site. She is a below-grade level student, but she really stuck with one game where she had to measure and weigh. It was good for her."

Perceptions of Cyberchase: Perhaps because of the perceived learning discussed above, virtually all of the teachers were quite positive about the educational value and usefulness of *Cyberchase*, and about its appeal for their students. Examples of teacher comments included:

"The characters in the video definitely kept students engaged. As a teacher I believe the hardest part is keeping students interested in the topic."

"Several of the episodes and games were great and matched the grade level content standards that I am required to cover. They allowed students to experience the same content in new and interesting ways."

"I like it when my kids like things that are good for them. Not like spinach. My kids attitudes are great. And Cyberchase helped."

"I've learned some new things too. That never hurts. ... I'm now spending more time on richer problems."

Appeal was strong among children, too. When to rate the appeal of *Cyberchase* (whichever combination of materials they had used) on a five-point scale of "Great-Good-OK-Not So Good-Terrible," 87% of children rated *Cyberchase* as either "good" or "great." No one rated it below "OK."

Interestingly, although the educational effects of *Cyberchase* appeared to be driven more by the television series than by the online games (in that the Web Only group often showed smaller effects than the other *Cyberchase* groups) – and although all of the groups rated the appeal of *Cyberchase* positively – appeal was stronger among groups that used the Web site either exclusively or alongside other materials ($F_{2,151} = 3.78$, $p < .05$).

CONCLUSIONS AND IMPLICATIONS

Together, these data indicate that children's naturalistic use of multiple media often spans multiple media platforms, and that there are indeed benefits to learning from multiple media over learning from a single medium. Moreover, the data also suggest possible ways in which these benefits might arise.

Use of Multiple Media

Obviously, the benefits of cross-platform learning can arise only if children choose to use multiple media platforms in the first place. Fortunately, the results of the present study suggest that they do. Data from the naturalistic phase clearly indicate that children's use of *Cyberchase* was consistent over time and spanned multiple media. Those children who chose to use *Cyberchase* typically did not engage in one-time use. Instead, they became "*Cyberchase* fans" whose interest in *Cyberchase* sustained itself over a period of several months and carried over to both television and the Web.

Certainly, we must be careful in attempting to generalize from *Cyberchase* to children's use of other media-based STEM projects. Every project is different and may be used differently, either by design or because of the interests of the children who use them. However, there are good reasons to believe that the patterns of use observed for *Cyberchase* may be typical of children's media use as a whole. First, children's reports regarding their use of *SpongeBob Squarepants* during the naturalistic phase followed similar patterns. Like *Cyberchase*, use of *SpongeBob Squarepants* was consistent across the naturalistic phase, and significant relationships were found between use of the *SpongeBob Squarepants* television series and Web site. Second, recent research on children's Web use (unrelated to *Cyberchase*) also supports the relationship between use of television and the Web. Data on Web usage in 2009 and 2010 found that approximately one-half of the 10 most popular Web sites for children were associated with television programs and characters such as those found on Nickelodeon, Cartoon Network, or PBS (e.g., Kido'z, 2009; Nielsen Online, 2010).

In the present study, as in past research, use of *Cyberchase* was not driven by either gender, ethnicity, mathematics ability, or prior interest in mathematics as a favorite subject. Rather, past research found that use was motivated by *Cyberchase*'s entertainment value – its characters, stories, etc. (Fisch, 2005). Since children in the present study rated the appeal of *Cyberchase* highly, that appears to have been the case here as well. Thus, it appears that appealing STEM media have the potential to attract a diverse audience (including children who do not have a prior affinity for the relevant educational content), and to encourage them to continue their engagement over time and across media.

Benefits for Learning

Past research has shown that children's knowledge of mathematics and problem-solving skills can be enhanced through use of educational television and computer games (e.g., Clements, 2002; Fisch, 2004), and, more specifically, by the *Cyberchase* television series (Fisch, 2003; Rockman Et Al., 2002). The present data replicate this finding for the television series and extend it to other *Cyberchase* media as well. While playing *Cyberchase* games online, children engaged in increasingly sophisticated cycles of problem solving that resembled the progression that past research has documented for classroom learning (e.g., Lesh, 2000). Subsequently, users of *Cyberchase* media demonstrated significantly greater gains in problems-solving performance than non-users. Interestingly, these gains were not manifest in children's simply doing a greater number of things while working on the tasks, but rather in their using a greater *variety* of strategies and heuristics, and in using those strategies and heuristics more effectively. These sorts of quantitative results fit well with our qualitative observations of children demonstrating top-down planning and persistence in the posttest.

As noted earlier, many of the pretest-posttest effects on problem solving appeared to be driven more by the *Cyberchase* television series than by the Web site. Perhaps this is to be expected, since television is the central component of the *Cyberchase* project. Not only did the children in the present study spend more time with the television series than the Web site (because our experimental treatment was designed to simulate real-world use, in which the television series is used more often than the Web site), but the *Cyberchase* television series carries far more explanation of embedded math concepts and problem solving than the online games do. The story-based format of the *Cyberchase* television series also may have played a role, by presenting models of successful problem solving in the context of compelling stories. Indeed, numerous researchers have pointed to the power of narrative in conveying and representing information (e.g., László, 2008; Schank & Abelson, 1995). As one of our participating teachers put it, *"I think kids remember things longer when they associate things with stories. They sometimes say, 'This is like when [the Cyberchase characters did something]."*

Because all of the instructional materials in this study were taken from *Cyberchase*, the present data should not be taken as evidence that television necessarily has greater educational potential than interactive games. It is quite possible that another project, designed with online games as its centerpiece, might find stronger effects for its games than for its video component. The present study was not intended to place different media on equal footing to discover which one is "best," but rather, to explore how different media might interact to yield cumulative effects, as we shall now discuss.

Cross-Platform Learning

One of the primary questions addressed by this study was how learning from multiple media compares to learning from a single medium, particularly with regard to transfer of learning. Our significant effects on problem solving suggest that children did engage in transfer of learning, applying the skills and concepts modeled in *Cyberchase* to new problems encountered in the posttest. Indeed, approximately one-third of the children who used *Cyberchase* spontaneously recalled information from *Cyberchase* explicitly while working on one of the posttest tasks, making the source of their inspiration apparent.

Such transfer appeared to occur more among children who used multiple media. In the pretest-posttest problem-solving tasks, many of the observed effects were stronger among the DVD + Web group than among either the DVD Only or (especially) the Web Only group. Contrary to our expectations, the same was not true of the All Materials group, which used all of the same materials as the DVD + Web group plus teacher-led hands-on materials. We cannot be certain why the All Materials group did not perform at the same level as the DVD + Web group, but we hypothesize that it may be because the All Materials group was the only one that used *Cyberchase* materials every day; perhaps this schedule was excessive in light of all of the other constraints on teachers' schedules, and was simply too much for participants to integrate effectively. Further research would be necessary to determine whether an "optimal level" of media use exists. Nevertheless, data from the DVD + Web group suggest that cross-platform learning can hold benefits for transfer of learning.

The benefits of cross-platform learning in promoting transfer were even more apparent in our online tracking data. These data revealed that children who used multiple media employed more sophisticated strategies while playing three online games, and produced more correct responses while playing two of the three games. Just as in the posttest tasks, it appears that children took the educational content they encountered in one medium (television and/or hands-on activities) and applied it while engaging with math content in another medium (online games). This transfer of learning supported their interaction with the second medium, allowing children to apply more sophisticated approaches and producing a richer, more successful engagement with the material.

Why, then, did cross-platform learning contribute toward transfer of learning – and toward greater transfer to posttest tasks? One possible explanation is simply that children who used multiple *Cyberchase* media spent more time engaging with their embedded mathematics content. To some degree, this explanation is probably at least partially correct. Indeed, one of the chief purposes of informal education is precisely that -- to encourage children to spend more time with educational content than they would otherwise. However, time clearly cannot explain the present findings by itself, because the benefits found for the DVD + Web group were not equaled by the All Materials group, which devoted even more time to *Cyberchase* activities. If time were the sole explanation, the gains shown by the All Materials group would have been at least as large as those of the DVD + Web group, if not larger.

A more promising explanation may lie in the concept of *varied practice* discussed in the educational research literature on transfer of learning (e.g., Gick & Holyoak, 1983; Salomon & Perkins, 1989; Singley & Anderson, 1989). In varied practice, learners are provided with multiple examples of the same concept or repeated practice of a skill in multiple contexts, which increases the likelihood that the learner will apply the material in new tasks or situations as well. As children in the present study encountered mathematics and problem-solving content in multiple *Cyberchase* media, they were clearly engaged in varied practice, especially in those instances where there was close alignment among the content of a related television episode, hands-on activity, and online game. Effects within the online tracking data attest to children's connecting the content of the different media, and even applying the content learned from one medium *while they were learning from the other*. Not only did children gain additional, varied practice by using multiple media, but their engagement with the latter medium was richer and more sophisticated as well. In this way, cross-platform learning has the potential to support learning by contributing to two types of transfer: transfer across educational media platforms (resulting in richer engagement and understanding), and transfer from educational media to new problems or situations encountered subsequently (such as our posttest assessments).

Moreover, it is quite possible that transfer may even be *facilitated* by the presence of the same characters and contexts across media. Past research on transfer of learning has shown that transfer is more likely to occur when two situations appear similar on their face (*surface structure similarity*) than when they are dissimilar on their surface but rest on similar underlying principles (*deep structure similarity*; e.g., Bassok & Holyoak, 1993; Gentner & Forbus, 1991). Thus, for example, encountering *Cyberchase* characters in an online game might lead children to think of other times when they saw the same characters (e.g., on television). This could facilitate the transfer of information and skills from one medium to another, in a way that seeing different characters on television and in a game might not.

Benefits for Attitude

Apart from its potential benefits for transfer of learning, the consistency of characters and contexts across media might contribute to attitudinal effects as well. Under Hidi and Renninger's (2006) four-stage model of interest development, interest in a subject such as mathematics originates as interest sparked by the context in which the math is embedded (*triggered situational interest*). Subsequently, interest can be maintained over a longer period by the context (*maintained situational interest*), after which it may evolve into interest in the mathematics itself (*emerging individual interest* and *well-developed individual interest*).

This model fits the present data well. Not only did we find significant associations between naturalistic use of the *Cyberchase* television series and Web site over time, but we also found that children who used multiple media were significantly more motivated to continue playing online *Cyberchase* games. The appeal of *Cyberchase* (as seen in children's appeal ratings at the end of the study) appears to have motivated children's continued use of *Cyberchase*, both over time and across media platforms (as seen in the naturalistic phase).

In Hidi and Renninger's terms, triggered situational interest contributes to maintained situational interest, with the potential to develop into individual interest. Or, in layman's terms, when children become fans of *Cyberchase* and spend more time with various *Cyberchase* media, they are spending more time engaged in substantive, enjoyable informal mathematics activities. Such activities have the potential to contribute to emerging interest in mathematics, as well as to learning.

Implications for the Design of ISE Media

When designing future multiple-media projects for informal education, the present data suggest that it is not merely the case that "more" is always better. As noted earlier, children in the DVD + Web group showed more consistent gains than groups that used only one medium – but they also often showed greater gains than children who used all three types of *Cyberchase* media. (Interestingly, another recent study also found that the strongest effects were not always found among the experimental group that used the greatest amount of media [Fisch et al., submitted for publication].) Further research is needed to determine whether we are correct in hypothesizing that there may be an optimal level of media use in the classroom, beyond which teachers (or even children) find the media less useful – and, if so, what that level might be.

Beyond simply the amount of media used, the data also suggest ways in which media can be designed to maximize their educational power:

- *Explanation and scaffolding:* We believe that one reason why effects were often driven more by the *Cyberchase* TV series than the online games may be that the television series provided more explanation of the relevant mathematical concepts as they used characters and narrative to model successful problem solving. If so, this argues for the need for educational media (in any medium) to provide, not only opportunities for children to exercise their existing and emerging skills, but also explanatory support and scaffolding when needed.

- *Narrative:* Researchers such as Schank and Abelson (1995) have theorized that narrative can serve as a powerful means for conveying information, and for organizing and storing information in memory. The present data are consistent with this view, in that pretest-posttest effects were often strongest among children who viewed the *Cyberchase* television series, and qualitative observations revealed instances of children explicitly referring to *Cyberchase* stories and characters as they worked on problems in the posttest. This is not to say, of course, that non-narrative formats (e.g., games, live demonstrations) cannot also convey educational content effectively. However, our findings speak to the potential for narrative to play an important role in mathematics education.

- *Complementary media:* To facilitate the sorts of transfer of learning and attitudinal effects discussed earlier, consistent characters and contexts, as well as complementary educational content, should be employed across media. In the case of *Cyberchase*, narrative media, such as video, supply explanation of content and models of successful problem solving, whereas participatory (interactive and hands-on) media that provide opportunities for children to exercise these skills themselves. The use of a common world and characters can encourage linkages of content from one medium to another, while the appeal of children's experience in one medium can enhance their motivation to engage with other educational media that employ the same characters.

- *Convergent media:* Together, the above points suggest intriguing possibilities for convergent media, in which the narrative and explanatory power of video, the participatory strength of interactive games, and the potential for scaffolding inherent in adult-mediated hands-on media can be combined in a single media-based experience. For example, consider an interactive game in which the "hint" button pulls up an explanatory video clip, or imagine a video with an embedded interactive game that allows the viewer to use mathematics to help the protagonist achieve her goal in the video.

In these ways, we can build on the lessons learned from past and current research, both to stimulate future research and – even more importantly – to build projects that will take even better advantage of the power of educational media to help children learn.

REFERENCES

Bassok, M., & Holyoak, K.J. (1993). Pragmatic knowledge and conceptual structure: Determinants of transfer between quantitative domains. In D.K. Detterman & R.J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction*. Norwood, NJ: Ablex.

Begel, E. (1979). *Critical variables in mathematics education: Findings from a survey of the empirical literature.* Washington, DC: Mathematics Association of America.

Bransford, J.D., Brown, A.L., & Cocking, R.R. (Eds.). (1999). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.

Clements, D.H. (2002). Computers in early childhood mathematics. *Contemporary Issues in Early Childhood, 3,* 2, 160-181.

Crane, V. , Nicholson, H.J., Chen, M., & Bitgood, S. (Eds., 1994). *Informal science learning: What the research says about television, science museums, and community-based projects*. Ephrata, PA: Science Press.

Debold, E., Hall, E.R., Fisch, S.M., Bennett, D.T., & Solan, S.V. (1990). *Children's problem-solving behavior and their attitudes toward mathematics: A study of the effects of Square One TV (Vol. 3: Children's attitudes toward mathematics and the effects of Square One TV)*. New York: Children's Television Workshop.

Dweck, C.S., & Elliott, E.S. (1983). Achievement motivation. In P.H. Mussen & E.M. Hetherington, *Handbook of child psychology, vol. 4* (pp. 643-691). New York: Wiley.

Falk, J.H., Donovan, E., & Woods, R. (Eds., 2001). *Free-choice science education: How we learn science outside of school (Ways of knowing in science and mathematics)*. New York: Teachers College Press.

Fay, A.L., Teasley, S.D., Cheng, B.H., Bachman, K.M., & Schnakenberg, J.H. (1995). *Children's interest in and their understanding of science and technology: A study of the effects of Cro*. Pittsburgh, PA: University of Pittsburgh and New York: Children's Television Workshop.

Fennema, E. & Sherman, J.A. (1976). Fennema-Sherman mathematics attitude scales. Instruments designed to measure attitudes toward the learning of mathematics by males and females. *JSAS Catalog of Selected Documents of Psychology, 6,* 31. (Ms. No. 1225)

Fisch, S.M. (2003). *The impact of Cyberchase on children's mathematical problem solving: Cyberchase season 2 summative study*. Teaneck, NJ: MediaKidz Research & Consulting.

Fisch, S.M. (2004). *Children's learning from educational television: Sesame Street and beyond*. Mahwah, NJ: Lawrence Erlbaum Associates.

Fisch, S.M. (2005). *Cyberchase parent Web survey*. Teaneck, NJ: MediaKidz Research & Consulting.

Fisch, S.M. (2006). *Cyberchase* season 5 pilot summative research: How are outreach materials used? Teaneck, NJ: MediaKidz Research & Consulting.

Fisch, S.M. (2007). *Season 5 pilot summative research: Quasi-experimental pilot test*. Teaneck, NJ: MediaKidz Research & Consulting.

Fisch, S.M., Hsueh, Y., Zhou, Z., Xu, C.J., Hamed, M., Khader, Z., Langsten, R., Noriega, G.M., Cespedes, A.H., Druin, A., & Guha, M.L. (submitted for publication). Can children learn global citizenship from educational media?: An international assessment.

Fisch, S.M., Lesh, R., Motoki, E., Crespo, S., & Melfi, V. (in press). Children's mathematical reasoning in online games: Can data mining reveal strategic thinking? *Child Development Perspectives*.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155-170.

Gentner, D., & Forbus, K.D. (1991). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science, 19*, 141-205.

Gick, M.L., & Holyoak, K.J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*, 1-38.

Greeno, J.G. Moore, J.L., & Smith, D.R. (1993). Transfer of situated knowledge. In D.K. Detterman & R.J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp. 99-167). Norwood, NJ: Ablex.

Hall, E.R., Esty, E.T., & Fisch, S.M. (1990). Television and children's problem-solving behavior: A synopsis of an evaluation of the effects of *Square One TV*. *Journal of Mathematical Behavior, 9*, 161-174.

Hannula, M.S. (2002). Attitude towards mathematics: Emotions, expectations, and values. *Educational Studies in Mathematics, 49*, 25-46.

Haskell, R.E. (2001). *Transfer of learning: Cognition, instruction, and reasoning*. San Diego: Academic Press.

Hidi, S., & Renninger, K.A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*, 111-127.

Hoffman, L. , Krapp, A., & Renninger, K.A. (Eds.; 1998), *Interest and learning: Proceedings of the Seeon conference on interest and gender* (pp. 197-204). Kiel, Germany: IPN.

Holyoak, K.J. (1985). The pragmatics of analogical transfer. In G.H. Bower (Ed.), *The psychology of learning and motivation* (pp. 59-87). New York: Academic Press.

Kelly, A.E., & Lesh, R.A. (2000). *Handbook of research design in mathematics and science education*. Mahwah, NJ: Lawrence Erlbaum Associates.

Kulm, G. (1980). Research on mathematics attitude. In Shumway, R.J. (Ed), *Research in mathematics education* (pp. 336-387). Reston, VA: National Council of Teachers of Mathematics.

Kunkel, D. (2001). Children and television advertising. In D. G. Singer & J. L. Singer (Eds.), *Handbook of Children and the Media* (pp. 375-393). Thousand Oaks, CA: Sage Publications.

Kido'z. (2009). 10 most popular Websites for kids. Available online: http://kidoz.net/blog/10-most-popular-websites-for-kids/.

László, J. (2008). *The science of stories: An introduction to narrative psychology*. New York: Routledge.

Lesh, R. & Doerr, H. (2002) *Beyond constructivist: A models and modeling perspective on mathematics teaching, learning, and problems solving*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Lesh, R., Hamilton, E., & Kaput, J. (Eds.; 2007). *Models and modeling as foundations for the future in mathematics education.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Lesh, R.A., Hoover, M., Hole, B., Kelly, A., & Post, T. (2000). Principles for developing thought-revealing activities for students and teachers. In A.E. Kelly & R.A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 591-646). Mahwah, NJ: Lawrence Erlbaum Associates.

Lester, F.K. (Ed.; 2007). *Second handbook of research on mathematics teaching and learning.* Reston, VA: National Council of Teachers of Mathematics & Information Age Publishing.

National Council of Teachers of Mathematics (1993). *Assessment standards for school mathematics*. Reston, VA: Author.

Nicholson, H.J., Hamm, J.K., & Weiss, F.L. (1991). *Operation SMART research tool kit: A package of program evaluation activities for girls*. Indianapolis, IN: Girls Inc.

Nielsen Online. (2010). The most popular Websites for kids, teenagers, and young adults. Available online: http://www.nielsen-online.com/pr/pr_080815_UK.pdf.

Petty, R. E., Priester, J. R., & Briñol, P. (2002). Mass media attitude change: Implications of the elaboration likelihood model of persuasion. In J. Bryant & D. Zillmann, (Eds.), *Media* *effects: Advances in theory and research* (2nd ed., pp. 155-199). Hillsdale, NJ: Lawrence Erlbaum Associates.

Pintrich, P.R., & Schunk, D.H. (2002). *Motivation in education: Theory, research, and applications* (2nd ed.). Upper Saddle River, NJ: Merrill/Prentice-Hall.

Polya, G. (1957). *How to solve it: A new aspect of mathematical method*. New York: Doubleday Anchor Books.

Research Communications Ltd. (1992). *The impact of using the FUTURES series on junior high school students*. Dedham, MA: Author.

Rideout, V.J., Foehr, U.G., & Roberts, D.F. (2010). *Generation M^2: Media in the lives of 8- to 18-year-olds.* Menlo Park, CA: Kaiser Family Foundation.

Rockman Et Al. (2002). *Evaluation of Cyberchase phase one pilot study: Vol. 1, executive summary*. San Francisco, CA: Author.

Salomon, G., & Perkins, D.N. (1989). Rocky roads to transfer: Rethinking mechanisms of a neglected phenomenon. *Educational Psychologist, 24*, 2, 113-142.

Schank, R.C., & Abelson, R.P. (1995). Knowledge and memory: The real story. In R.S. Wyer (Eds.), *Knowledge and memory: The real story* (pp. 1-85). Hillsdale, NJ: Lawrence Erlbaum Associates.

Schoenfeld, A. (1992). Learning to think mathematically: problem solving, metacognition, and sense making in mathematics. In D.A. Grouws (Ed.), *Handbook of research on mathematics teaching* (pp. 334-370). New York: MacMillan.

Schwartz, D. L., Bransford, J. D., & Sears, D. L. (2005). Efficiency and innovation in transfer. In J. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 1 - 51). CT: Information Age Publishing.

Siegler, R.S. (2007). Cognitive variability. *Developmental Science, 10*, 104-109.

Singley, M.K., & Anderson, J.R. (1989). *The transfer of cognitive skill*. Cambridge, MA: Harvard University Press.

U.S. Department of Education Office of Innovation and Improvement. (2010). Ready-to-learn television program: Notice inviting applications for new awards for fiscal year (FY) 2010. *Federal Register, 75,* 13515- 13521.

Appendices:

Children's Learning from Multiple Media in Informal Mathematics Education

Shalom M. Fisch
MediaKidz Research & Consulting

Richard Lesh, Elizabeth Motoki
Indiana University

Sandra Crespo, Vincent Melfi
Michigan State University

TABLE OF CONTENTS

APPENDIX A
Background questionnaire and sample *Cyberchase* journal

Background Measure

Name: ________________________________

Grade: ________________________________

Class: _________________________________

Boy/Girl: _______________________________

Date I was born: __________________________

My favorite subject in school is: _____________

I have seen *Liberty's Kids* on TV:

- ___ Never (0 times)
- ___ A few times (1-5 times)
- ___ A lot (6-10 times)
- ___ A whole lot (more than 10 times)

I have seen *SpongeBob Squarepants* on TV:

- ___ Never (0 times)
- ___ A few times (1-5 times)
- ___ A lot (6-10 times)
- ___ A whole lot (more than 10 times)

I have seen *Scooby Doo* on TV:

- ___ Never (0 times)
- ___ A few times (1-5 times)
- ___ A lot (6-10 times)
- ___ A whole lot (more than 10 times)

I have seen *Cyberchase* on TV:

- ___ Never (0 times)
- ___ A few times (1-5 times)
- ___ A lot (6-10 times)
- ___ A whole lot (more than 10 times)

I have been to the Cartoon Network Web site:

- ___ Never (0 times)
- ___ A few times (1-5 times)
- ___ A lot (6-10 times)
- ___ A whole lot (more than 10 times)

I have been to the *Cyberchase* Web site:

- ___ Never (0 times)
- ___ A few times (1-5 times)
- ___ A lot (6-10 times)
- ___ A whole lot (more than 10 times)

I have been to the *Liberty's Kids* Web site:

- ___ Never (0 times)
- ___ A few times (1-5 times)
- ___ A lot (6-10 times)
- ___ A whole lot (more than 10 times)

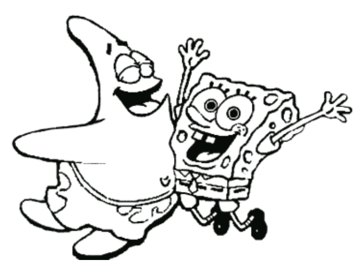
I have been to the Nickelodeon Web site:

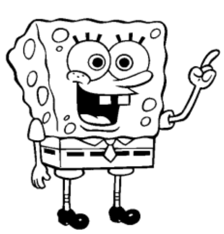
- ___ Never (0 times)
- ___ A few times (1-5 times)
- ___ A lot (6-10 times)
- ___ A whole lot (more than 10 times)

Sample *Cyberchase* journal pages

APRIL

My Cyberchase and SpongeBob Journal





My Name: ________________________________

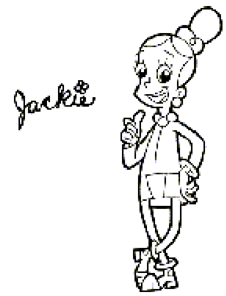
My Class: ________________________________

My Grade (circle one): 3rd 4th

I am a (circle one): Girl Boy

My Cyberchase login: ________________________

My password: ______________________________



WEEK 1 (March 30-April 5)

CYBERCHASE

1) Did you watch *Cyberchase* on TV this week? Circle your answer:

Yes No

2) If you watched *Cyberchase* this week, please circle the days when you watched it:

Sunday Monday Tuesday Wednesday Thursday Friday Saturday

3) Did you visit the *Cyberchase* Web site this week? Circle your answer:

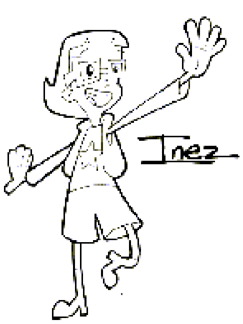
Yes No

4) If you visited the *Cyberchase* Web site this week, please circle the days when you visited it:

Sunday Monday Tuesday Wednesday Thursday Friday Saturday

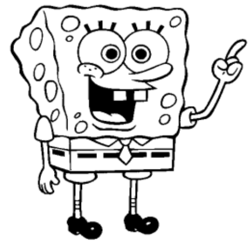
5) If you visited the *Cyberchase* Web site, how much time did you spend there?

6) If you visited the *Cyberchase* Web site, what did you do there? (Which games did you play, etc.?)



WEEK 1 (March 30-April 5)

SPONGEBOB SQUAREPANTS



1) Did you watch *SpongeBob* on TV this week? Circle your answer:

Yes No

2) If you watched *SpongeBob* this week, please circle the days when you watched it:

Sunday Monday Tuesday Wednesday Thursday Friday Saturday

3) Did you visit the *SpongeBob* Web site this week? Circle your answer:

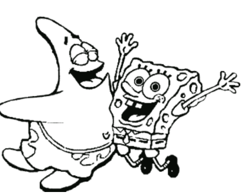
Yes No

4) If you visited the *SpongeBob* Web site this week, please circle the days when you visited it:

Sunday Monday Tuesday Wednesday Thursday Friday Saturday

5) If you visited the *SpongeBob* Web site, how much time did you spend there?

6) If you visited the *SpongeBob* Web site, what did you do there? (Which games did you play, etc.?)



Cross-

APPENDIX B

Treatment schedule

Experimental phase: Treatment

The following table lists the *Cyberchase* materials used in the experimental treatment. The DVD Only group watched the TV episodes listed in the "Show" column. The Web Only group played the games listed in the "Web game" column. The DVD + Web group did both types of activities. And the All Materials group did all of the activities listed in the table.

Week	Day	Show (Content)	Web game	Outreach activity	Theme
1	1	Team Spirit (Planning based on past performance)	Cyberolympics (planning based on past performance)		Org. data
	2	Escape from Merlin's Maze (Builders' math – length of lever)			Meas.
	3	Past Perfect Prediction (Predicting from data)		Play Slugball (Sports; predicting from data)	Org. data
2	1	Of All the Luck (Logic/Venn diagram)	Logic Zoo (Venn diagram)	Sort it Out (Venn diagram)	Org. data
	2	Sensible Flats (Area/irregular shapes)			Meas.
	3	R-Fair City (Probability & chance)			Org. data
3	1	Return to Sensible Flats (Line graphs)		Fill 'Er Up (graphing)	Org. data
	2	Time to Cook (Elapsed time)			Meas.
	3	Castleblanca (Meaning from data)	Railroad Repair (decimals & planning – note: off-topic, but strong problem solving)		Org. data
4	1	A Day at the Spa (Combinatorics)		Cyber Chow Combos (combinatorics)	Org. data
	2	Size Me Up (Scale and size)	Jigsaw Puzzle Size-Up (scale and size)		Meas.
	3	Perfect Score (Creating scoring system)			Org. data

5	1	Totally Rad (Area/perimeter)	Cyberchase Airlines Builder (area/perimeter)	Skate Borders (area/perimeter)	Meas.
	2	Starlight Night (Finding a simpler case)			Org. data
	3	A Broom of One's Own (Time/distance/speed)			Meas.
6	1	Measure for Measure (Choosing units of measure for volume)	Pour to Score (liquid volume)		Meas.
	2	Snow Day to Be Exact (Estimation)		Pool Party (estimation)	Meas.
	3	Whale of a Tale (Ballpark estimation)			Meas.
7	1	Ecohaven CSE (Body math)	Sleuths on the Loose (Body math)	Footprint Files (Body math)	Meas.
	2	Step By Step (Builders' math – length of bridge)			Meas.
	3	A Change of Art (Interpreting change in line graphs)			Org. data
8	1	Unhappily Ever After (Builders' math – cut to fit)	U Fix It (cut to fit)	Put a Lid on It (Builder's math; carpenter's square)	Meas.
	2	Fortress of Attitude (Linear measurement)			Meas.
	3	The Poddleville Case (Patterns)			Org. data

Liberty's Kids episodes for No Exposure (control) classrooms

The following table lists the episodes of the American History series *Liberty's Kids* that the No Exposure group watched during the weeks when other groups used *Cyberchase* materials.

Week	Day	Show
1	1	The Boston Tea Party (disc 1, episode 1)
	2	The Intolerable Acts (disc 1, episode 2)
	3	United We Stand (disc 1, episode 3)
2	1	Liberty or Death (disc 1, episode 4)
	2	Midnight Ride (disc 1, episode 5)
	3	The Shot Heard Round the World (disc 1, episode 6)
3	1	Green Mountain Boys (disc 1, episode 7)
	2	The Second Continental Congress (disc 2, episode 1)
	3	Bunker Hill (disc 2, episode 2)
4	1	Washington Takes Command (disc 2, episode 3)
	2	Postmaster General Franklin (disc 2, episode 4)
	3	Common Sense (disc 2, episode 5)
5	1	The First Fourth of July (disc 2, episode 6)
	2	New York, New York (disc 2, episode 7)
	3	The Turtle (disc 3, episode 1)
6	1	One Life to Lose (disc 3, episode 2)
	2	Captain Molly (disc 3, episode 3)
	3	American Crisis (disc 3, episode 4)
7	1	Across the Delaware (disc 3, episode 5)
	2	An American in Paris (disc 3, episode 6)
	3	Sybil Ludington (disc 4, episode 1)
8	1	Lafayette Arrives (disc 4, episode 2)
	2	The Hessians Are Coming (disc 4, episode 3)
	3	Valley Forge (disc 4, episode 4)

APPENDIX C
Paper-and-pencil problem-solving tasks with solution score coding schemes

Measurement Task: Roadmap (Experimental Pretest)

Two brothers, Alan and Zak, are arguing about which path is the longest way to walk home from school. There are four different paths that they can use. Can you use this map to show them which path is the longest? Please circle your answer:

A B C D

How do you know that path is the longest?

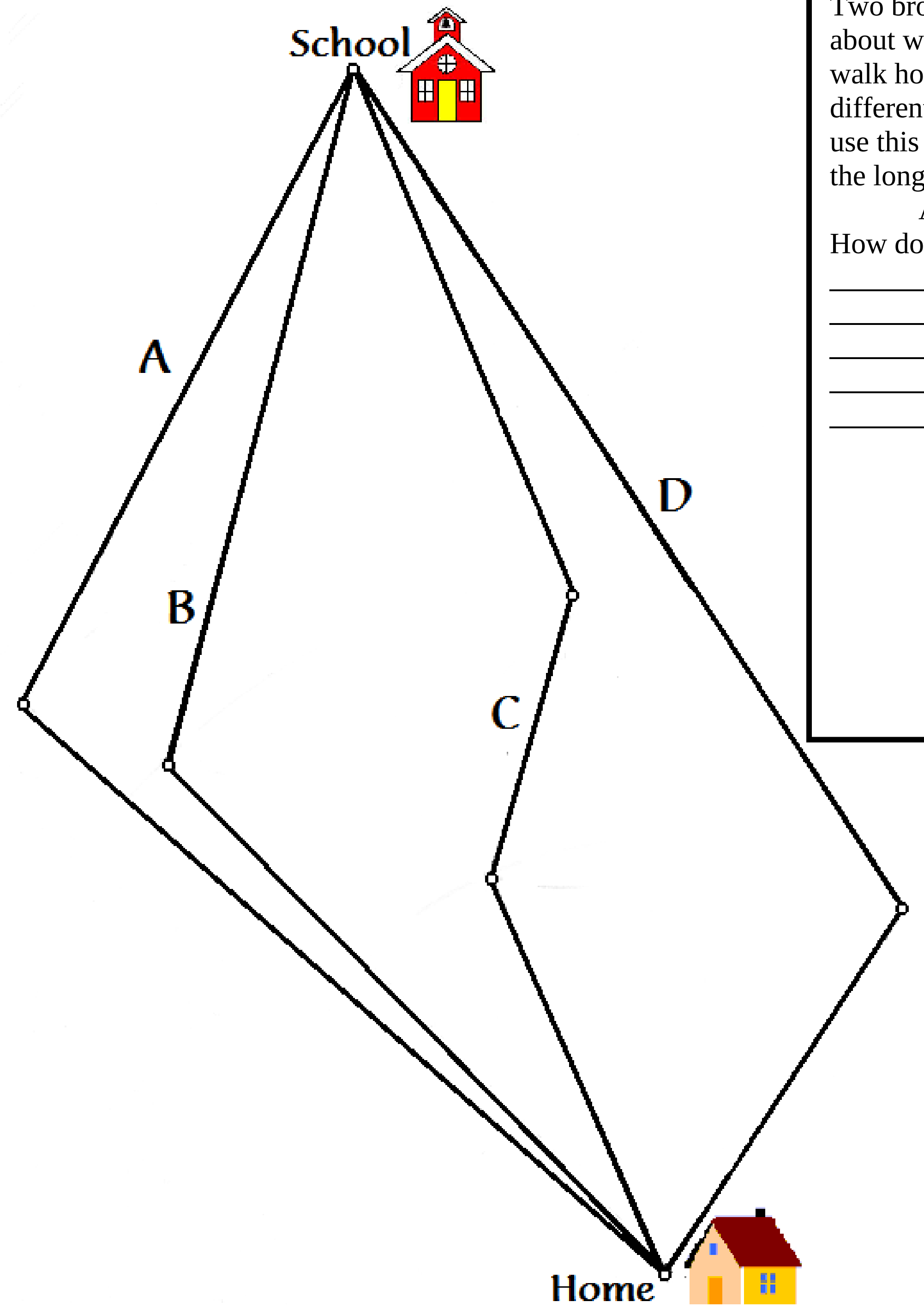
____

____

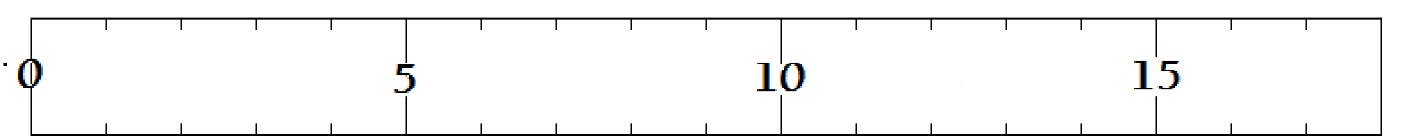
____

____

____



If you want, you can cut or tear out this ruler to help you.



Roadmap task (experimental pretest): Solution score coding scheme

There are at least two ways that this task might be approached – either by measuring the line segments and adding them up, or by some geometric argument. Accurate measurement will solve the task, but a purely geometric argument won't work. (A purely geometric argument could show that A is longer than B, but that's about all; it can't tell you anything about C and D.)

The following point scores are given cumulatively. Some recognition (1 point) is awarded for an attempt at a geometric argument, but mostly the points are awarded for successively better, more sophisticated, more complete, and more accurate measurement.

Note: The actual lengths of the paths are:

- A = 22 (10 + 12) – the longest path
- B = 20 (10 + 10)
- C = 18 (8 + 4 + 6)
- D = 20 (14 + 6)

The following is what I propose for various components of responses:

Add 0: for no answer at all OR **wrong road with vacuous reason**, e.g.:

B, "I look at B."

D, "because it looks the longest"

B, "It seems like it is the longest line"

C, "It starts on a longer spot"

Add 1: for choosing A (either by itself or with another path), regardless of anything else, with or without a "reason," e.g.:

A, "I know that A is the longest because every other one is shorter"

A, "I don't know"

A, "Because I looked at the picture. P.S. I cut out the ruler for fun."
(Note that this answer shows that merely cutting off the ruler is NOT an indication that measurement has been used!)

Add 1: for any geometric argument (regardless of which road the child says is longest), even though specious – e.g.:

B, "'Cause that road goes down to the middle and then cuts threw diagonally downward"

C, “It has the most turns.”

B, “Because B is the least bent.”

D, “Because you don’t have to turn you just go straight.”

D, “It is the longest because it is the furthest away.”

B, “It only turns once and looks long.”

B, “Because if you straighten it out it is the longest.”

B, “It can’t be C, because if you straighten it out it’s the same size as D + A.”

A, “It goes straight, and if you do D you will need to go straight and then go to the left.”

Add 1: for any evidence of measuring, even self-reported, with or without numbers, and regardless of accuracy – e.g.:

D, “D is the answer because I measured all of them and I found D was the answer.”

A, “I measured the lines and figured it out.”

B, “I know because I used the ruler and made it from start to finish the shortest way.”

A, “A because it is 20 inches and B is 19 inches and D is 18 inches and C is 17 inches.” (Note: These numbers are written next to the A, B, C, D but there is nothing indicating how the numbers were obtained. And they’re all incorrect.)

A, “I used the ruler.”

Add 1: for any evidence of measuring individual line segments and adding up those measurements, even if it’s not completely accurate, and even if the wrong route is chosen because of measurement or addition errors. The individual numbers for line segments don’t have to be given individually; if a child says that the path is more than 17 units long (which is the distance of a straight line from start to finish), we can assume that he/she added. For example, if a child says that B is 20 units long, then we can assume that B’s two line segments have been measured and added because there’s no other way to obtain 20. Examples:

B and C both circled, “I know because I measured.” Numbers are written for all the segments, and B and C both total 18 (B should be 20).

A and D are circled, “I used the ruler.” (A is labeled 22, B labeled 20; C is labeled 18; D is not labeled.)

B, (No reason given, but “10 + 11 = 21” is written; the actual lengths of the two segments of B are 10 and 10.)

Add 1: for accurate measurement of at least one path, regardless of whether C is chosen as the shortest. Examples:

B and C both circled, “I know because I measured.” Numbers are written for all the segments, and B and C both total 18 (B should be 20).

“A, Every corner I count what the number is then add it up until the dot. Then see which is longest,” with the numbers 22, 20, 18, and 20 written next to the letters in the box.

Add 1: for accurate measurements of all four possible paths, regardless of whether A is chosen as the longest. Note: This point is <u>in addition</u> to the 1 point that the child receives for measuring at least one path accurately. Thus, a child who measures one, two, or three paths correctly would receive 1 point for accurate measurement, whereas a child who measures all four paths correctly would receive 2 points for accurate measurement. Examples:

“A, Every corner I count what the number is then add it up until the dot. Then see which is longest,” with the numbers 22, 20, 18, and 20 written next to the letters in the box.

Examples of cumulative scores:

So, for example,

- A, “It goes straight, and if you do D you will need to go straight and then go to the left” would get:
  1 point for saying A is the longest, and
  1 point for a geometric reason

  for a total score of 2.

- “B,” (No reason given, but “10 + 11 = 21” is written; the actual lengths of the two segments of B are 10 and 10) would get:
  0 points for saying B is the shortest,
  1 point for evidence of measuring
  1 point for evidence of measuring individual line segments

  for a total score of 2.

- “A, Every corner I count what the number is then add it up until the dot. Then see which is longest,” with the numbers 22, 20, 18, and 20 written next to the letters in the box, would get:
  1 point for saying C is the shortest,
  1 point for evidence of measuring,

1 point for measuring line segments and adding,
1 point for accurate measurement of at least one path, and
1 point for accurate measurements of all four paths,

for a total score of 5.

Measurement Task: Roadmap (Experimental Posttest)

There are four trails on this map Amy can take to reach her family's cabin. Can you figure out which trail is the shortest? Please circle your answer:

A B C D

How do you know that trail is the shortest?

____

____

____

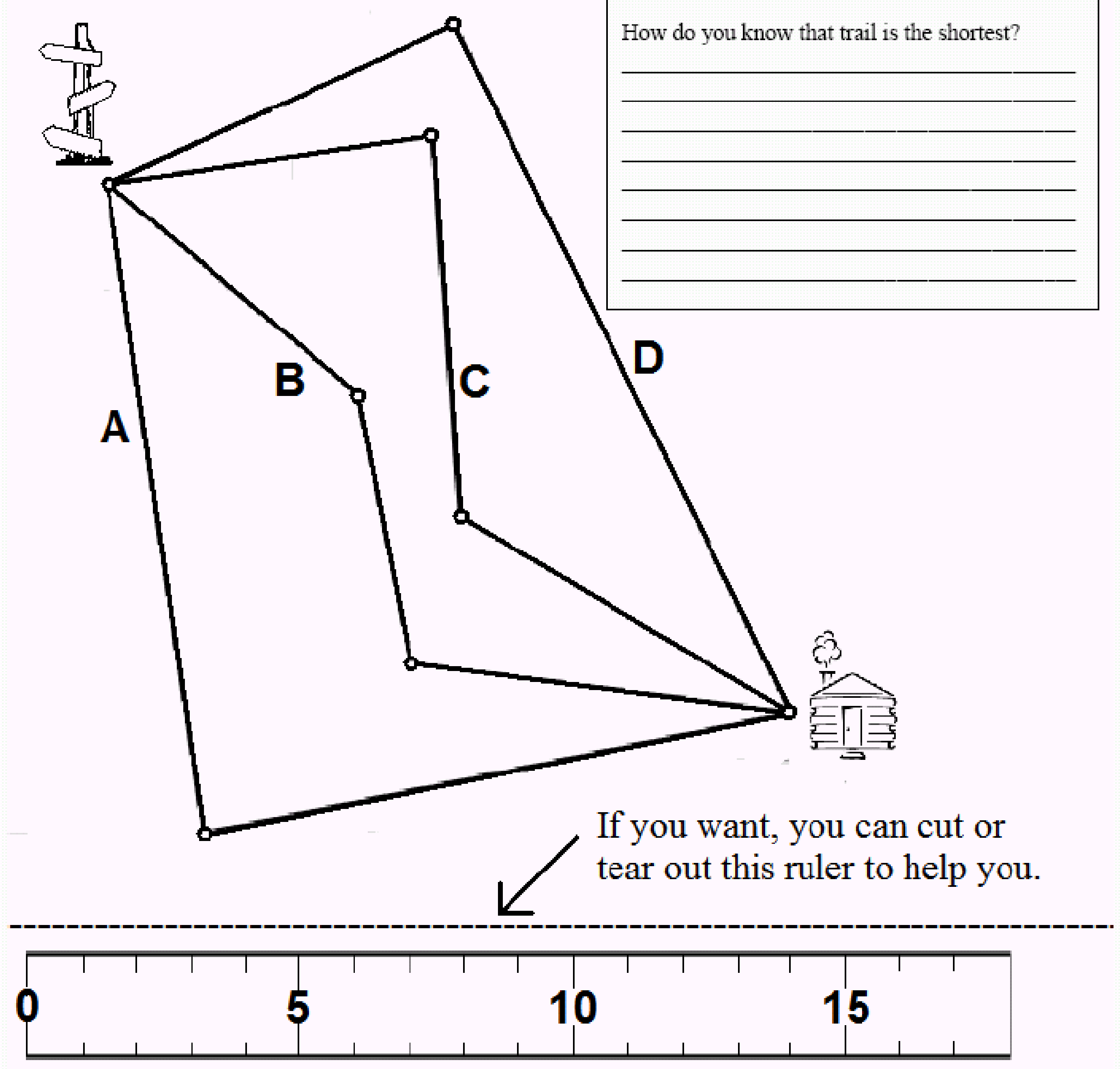
____

____

____

____

____



Roadmap task (experimental posttest): Solution score coding scheme

There are at least two ways that this task might be approached – either by measuring the line segments and adding them up, or by some geometric argument. Accurate measurement will solve the task, but a purely geometric argument is not sufficient.

The following point scores are given cumulatively. Some recognition (1 point) is awarded for an attempt at a geometric argument, but mostly the points are awarded for successively better, more sophisticated, more complete, and more accurate measurement.

Note: The actual lengths of the paths are:

- A = 23 (12 + 11)
- B = 18 (6+5+ 7) – the shortest path
- C = 20 (6 + 7 +7)
- D = 21 (7 + 14)

The following is what I propose for various components of responses:

Add 0: for no answer at all OR **wrong road with vacuous reason**, e.g.:

B, "I look at B."

A, "because it looks the shortest"

B, "It seems like it is the shortest line"

A, "It starts on a shorter spot"

Add 1: for choosing B (either by itself or with another path), regardless of anything else, with or without a "reason," e.g.:

C, "I know that B is the shortest because every other one is longer"

C, "I don't know"

C, "Because I looked at the picture. P.S. I cut out the ruler for fun"
(Note that this answer shows that merely cutting off the ruler is NOT an indication that measurement has been used!)

Add 1: for any geometric argument (regardless of which road the child says is shortest), even though specious – e.g.:

> B, "cause that road goes down to the middle and then cuts threw diagonally downward" [Ephraim L.]
>
> C, "He stats in the middle and C is one that's in the midale"
>
> A, "Becous you don't have to turn you just go strait"
>
> D, "Because it doesn't have a lot of turns"
>
> C, "You make lots of turns that make short"
>
> D, "it is the shortest because it is the closest"
>
> A, "it only turns once and looks short"
>
> B, "because if you straighten it out it is the shortest"
>
> B, "It can't be C b/c if you straighten it out it's the same size as D + A"
>
> B, "Because b is the least bent"
>
> C, "It goes strate and if you do A you will need to go down strate and then go to the righte"

Add 1: for any evidence of measuring, even self-reported, with or without numbers, and regardless of accuracy – e.g.:

> D, "D is the ansewer because I meshered all of them and I found D was the ansewer"
>
> C, "I measured the lines and figured it out"
>
> B, "I know because I used the ruler and made it from start to finish the shortest way"
>
> C, "C because it is 21 inches and A is 24 inches and D is 22 inches and B is 23 inches." (Note: These numbers are written next to the A, B, C, D but there is nothing indicating how the numbers were obtained. And they're all incorrect.)
>
> C, "I used the ruler."

Add 1: for any evidence of measuring individual line segments and adding up those measurements, even if it's not completely accurate, and even if the wrong route is chosen because of measurement or addition errors. The individual numbers for line segments don't have to be given individually; if a child says that the path is more than 16 units long (which is the distance of a straight line from start to finish), we can assume that he/she added. For example, if a child says that A is 23 units long, then we can assume that A's two line segments have been measured and added because there's no other way to obtain 23. Examples:

> B and C both circled, "I know because I measured." Numbers are written for all the segments, and B and C both total 20 (B should be 18)
>
> C and D are circled, "I used the ruler" (A is labeled 24, B labeled 23; C is labeled 21; D is not labeled.)
>
> B, (No reason given, but "6 + 5 + 7 = 23" is written; those are the lengths of the segments of B, if one includes part of the large dots.)

Add 1: for accurate measurement of at least one path, regardless of whether B is chosen as the shortest. Examples:

> B and C both circled, "I know because I measured." Numbers are written for all the segments, and B and C both total 20 (B should be 18)
>
> "C, Every corner I count what the number is then add it up until the dot. Then see which is shortest," with the numbers 23, 20, 18, and 21 written next to the letters in the box.

Add 1: for accurate measurements of all four possible paths, regardless of whether C is chosen as the shortest. Note: This point is <u>in addition</u> to the 1 point that the child receives for measuring at least one path accurately. Thus, a child who measures one, two, or three paths correctly would receive 1 point for accurate measurement, whereas a child who measures all four paths correctly would receive 2 points for accurate measurement. Examples:

> "C, Every corner I count what the number is then add it up until the dot. Then see which is shortest," with the numbers 23, 18, 20, and 21 written next to the letters in the box.

Examples of cumulative scores:

So, for example,

- "B, He stats in the middle and B is the one that's in the midale" would get:
  1 point for saying B is the shortest, and
  1 point for a geometric reason

  for a total score of 2.

- "C," (No reason given, but "6+7+7 = 20" is written; those are the lengths of the segments of C.) would get:
  0 points for saying C is the shortest,
  1 point for evidence of measuring

1 point for evidence of measuring individual line segments
1 point for evidence of accurate measuring at least one path

for a total score of 4.

- "B, Every corner I count what the number is then add it up until the dot. Then see which is shortest," with the numbers 23, 18, 20, and 21 written next to the letters in the box. would get:

  1 point for saying B is the shortest,
  1 point for evidence of measuring,
  1 point for measuring line segments and adding,
  1 point for accurate measurement of at least one path, and
  1 point for accurate measurements of all four paths,

  for a total score of 5.

Maximum total score is 5.

Organizing Data Task: Predicting from Data (Experimental Pretest)

Name: ______________________________ I am a (circle one): Girl Boy
My grade (circle one): 3 4 My teacher: __________________

It was summertime. The kids in a local children's summer theater were putting on a musical show, with performances on Wednesday and Saturday nights. Tickets were free but they weren't sure where the best places to hand them out would be. So, on Wednesday afternoon, they decided to go to different places around town, hand out tickets for that night's show, and see how many people took the tickets.

This table shows the places they handed out tickets, and how many people took the tickets.

Free Ticket Locations	**Number of Tickets**
Summer school entrance	16
Summer theater lobby	9
Library entrance	23
Grocery store entrance	44
Small movie theater lobby	12
Seven screen multiplex movie theater lobby	39

Can you help the kids figure out which locations work best? Please answer the following questions:

1. In which locations did they give out the most tickets?

____

____

2. In which locations did they give out the least tickets?

____

____

3. Why do you think they gave out more tickets in some locations than others?

____

____

____

4. Next, the kids need to hand out tickets for the performance on Saturday night. They can hand out tickets on Thursday, Friday, and Saturday, but there's a problem: Only one kid is available to hand out tickets on those days, so she can only go to one place. She has decided to hand out tickets at the grocery store, and wonders how many tickets she'll hand out.

Tell us what you think: About how many tickets will she probably hand out in the next few days? Why do you think so?

____

____

____

____

____

____

Predicting from data task (pretest): Solution score coding scheme

The following scores are to be awarded cumulatively.

PART 1

Add 0 for no answer at all, or for an answer to some other question, or for one of the three least effective locations, e.g.:

> "Small movie theater lobby"
> "44"

Add 1 point for identification of any of the top three locations; e.g.

> "Grocery store"
> "Multiplex theater"
> "Library"

Add 1 for identification of the top location ("Grocery store entrance"), with or without other effective locations. (Note that this is *in addition to* the previous point, so a child who lists the grocery store receives two points.) E.g.,

> "Grocery store"

"Grocery store entrance, multiplex theater, library"

So the highest possible score for Part 1 is 2 points.

PART 2

Part 2 is similar:

Add 0 for no answer at all, or for an answer to some other question, or for one of the three most effective locations.

Add 1 point for identification of any of the least effective three locations.
"Summer theater lobby"
"Small movie theater lobby"
"Summer school entrance"

Add 1 for identification the worst location ("Summer theater lobby"), even if other relatively ineffective locations are cited. (Again, this point is awarded *in addition to* the previous point.)

So highest possible score for Part 2 is 2 points.

PART 3

0 points for no answer, or completely irrelevant answers; e.g.

"Because it was a better solution."
"So it's easier."
"So they can watch a movie after they give out the tickets."

Add 1 point for that suggests any kind of comparison of the locations; e.g.

"Because lots of people go to the grocery store."
"Because [the multiplex] is bigger than the other movie theater."
"Because there were more people in some places and not as many in others."

Add 1 point for any suggestion that *more* people are reached in some locations than others; e.g.:

"Because more people go there."
"Because it's more public."
"Because not so many people go to summer school."
"Because some were a little and some were a lot."

So top possible score for Part 3 is 2 points.

PART 4
Note that there are two questions involved here – how many tickets? and why?

AMOUNT

0 points for no answer or irrelevant answer; e.g.

"They want to see the show."

Add 1 point if there is any amount given, even with no reason at all, and even if the quantity is not exactly specified or stated only as a comparison to some other number; e.g.:

"I think they will give out about 60 tickets."
"I think they will give out a lot of tickets because a lot of people go to the grocery store."
"200"
"40 more tickets because they will get more people."

Add 1 point if the quantity specified is 40 or greater (suggesting that the child recognizes that multiple days at the grocery store should elicit at least as much food as one day – with the potential for children to round 44 down to 40), e.g.:

"I think they will give out about 60"
"200."

Add 1 point if the amount given is 132, 88, or 44, which suggests that the child has performed a calculation to produce his or her estimate. (That's either 3 days or 2 days times the 44 tickets from one day at the grocery store.) OR if the amount is 120 (rounding the number of tickets down to 40 times 3 days).

Top score for this section of Part 4 is 3 points.

REASONS

Add 0 points for irrelevant reason, or no reason; e.g.
"So they see the show."
"So a lot of people come."
"I do not know why."

Add 1 point if there is any reference (overt or implied) to the original day at the grocery store or the 44 tickets distributed during the first day at the grocery store, e.g.

"About 30 tickets because some people saw the show already."
"They will give out 54 tickets because they're doing it for more days."
"I think about 44 tickets because that's how much last time."

Add 1 point if there is any reference to, or use of, the fact that tickets will be distributed for 3 days. (Note that this point would not be awarded to calculations that do not involve 3.) E.g.:

"About 132 because in one day it was 44." (i.e., implying that 44 x 3 = 132)

Add 1 point if there is any explicit calculation involving the number of ads and the number of pounds per ad, regardless of the number of days that is used, and even if the calculation is incorrect; e.g.

> "About 132 because in one day there was 44, and this is 3 days so 44 x 3 = 132." "I'm estimating if they give out the same amount each day like they did the first time, they'll give out 132 tickets." (44 x 3 is written on the side of the paper in the usual algorithm.)
> "About 130 tickets, because in one day they gave out 44 tickets, so 3 days would make about 130 tickets."
> "About 88 tickets because there will be more days." (A column addition of 44 + 44 = 88 is written.)

Add 1 point if there is any reference to "diminishing returns" or that people who come to the store more than once are not likely to take tickets every time; e.g.:

> "75, because some of the people might be the same and some might be different."
> "30, because some already went to the show on Wednesday."
> "About 100 because it's 3 days, but some of the people saw it already."

Top possible score for this section of Part 4 is 4 points.

So the top possible score for the entire Part 4 is 7 points.

Top possible score for the whole item is 13 points.

Organizing Data: Predicting from Data Task (Posttest)

Name: ________________________________ I am a (circle one): Girl Boy

My grade (circle one): 3 4 My teacher: ___________________

Kids are trying to raise money for the local animal shelter. Each day this week they have been asking shoppers at a Superette Food Store for donations. During the first week, here are the amounts they have collected (rounded to the nearest dollar):

Day of the week	Amount collected
Sunday	$49
Monday	$13
Tuesday	$15
Wednesday	$23
Thursday	$27
Friday	$74
Saturday	$68

Please answer the following questions:

1. On which days did they collect the most money?

____

____

2. On which days did they collect the least money?

____

____

3. Why do you think they got more money on some days than on other days?

____

____

____

____

____

4. The head of the Superette Food Store tells the kids that next week, **on just one day**, they can collect donations in front of all FIVE of his stores, all over the city.

Tell us what you think. Which day of the week should they pick, and how much money do you think they will get? Why do you think so?

____

____

____

____

____

____

____

____

Predicting from data task (posttest): Solution score coding scheme

The following scores are to be awarded cumulatively.

PART 1

Add 0 for no answer at all, or for an answer to some other question, or for one of the four least effective days, e.g.:

> "Tuesday"
> "23"

Add 1 point for identification of any of the top three days; e.g.

> "Friday"
> "Saturday"
> "Sunday"

Add 1 for identification of the top day ("Friday"), with or without other top days. (Note that this is *in addition to* the previous point, so a child who lists Friday receives two points.) E.g.,

> "Friday"

"Friday, Saturday, Sunday"

So the highest possible score for Part 1 is 2 points.

PART 2

Part 2 is similar:

Add 0 for no answer at all, or for an answer to some other question, or for one of the three most effective days.

Add 1 point for identification of any of the least effective four days.

"Monday"
"Tuesday"
"Wednesday"
"Thursday"

Add 1 for identification of the worst day ("Monday"), even if other relatively ineffective days cited. (Again, this point is awarded *in addition to* the previous point.)

So highest possible score for Part 2 is 2 points.

PART 3

0 points for no answer, or completely irrelevant answers; e.g.

"Because it was better."
"So it's easier."
"So they can play soccer afterwards."

Add 1 point for that suggests any kind of comparison of the days; e.g.

"Because lots of people go to the store on weekends."
"Because people are home on the weekend."
"Because people need extra food for parties on weekends."
"Because nobody shops on Monday."
"Because there were more people there on some days than others."
"Because more people had time on Saturday."

Add 1 point for any suggestion that *more* people are reached on some days than others; e.g.:

"Because more people go there on those days."
"Because it's more public."
"Because not so many people shop early in the week."
"Because some days were a little people and some were a lot."

So top possible score for Part 3 is 2 points.

PART 4

Note that there are three questions involved here – which day? how much money? and why?

DAY/AMOUNT

0 points for no answer or irrelevant answer; e.g.

"It's when they can go."

Add 1 point if any day or amount is given, even with no reason at all, and even if the quantity is not exactly specified or stated only as a comparison to some other number; e.g.:

"Thursday."
"I think they will get about $50."
"I think they will get a lot of money because a lot of people go to the grocery store that day."
"$200"

Add 1 point if Friday or Saturday is given, even with no reason at all, or **if the amount specified is $60 or greater** (suggesting that the child recognizes that multiple locations should elicit at least as much money as one on that same day – with the potential for children to round 68 down to 60 if they choose Saturday rather than Friday for some reason), e.g.:

"I think they will get about $80"
"$200."

(Note: a tighter scoring would require selecting Friday as the top day and $70 as the minimum value.).

Add 1 point if the amount given is 300, 340, 350, 370, or 450, which suggests that the child has performed a calculation to produce his or her estimate. (That's either $74 or a round down/up to $70 or $80 times 5 stores, or $68 or a round down/up to $60 or $70 times five stores).

(Note: a tighter scoring would require selecting Friday and amounts of 350, 370, or 450 as the amount given.)

Top score for this section of Part 4 is 3 points.

REASONS

Add 0 points for irrelevant reason, or no reason; e.g.
"So they get the most."

"So a lot of people come."
"I do not know why."

Add 1 point if there is any reference (overt or implied) to a particular day or to the amount collected, e.g.

"About $80 because they got most of that on Saturday."
"They will get $200 because they're going to more stores this time."
"I think Saturday is best for the money and kids can go then."

Add 1 point if there is any reference to, or use of, the fact that they will be collecting money at 5 locations. (Note that this point would not be awarded to calculations that do not involve 5.) E.g.:

"About $370 because at one place it was $74." (i.e., implying that 74 x5 = 370. Note that errors in calculation should not disqualify this entry if the multiplication can be discerned.)

Add 1 point if there is any explicit calculation involving the number of locations and the amount of money collected at each location, regardless of the numbers that are used, and even if the calculation is incorrect; e.g.

"About $370 because in one place there was $74, and this is 5 places so 74 x 3 = 370."
"I'm estimating they will get the same amount from each store like they did last time, so $370." (74 x 5 is written on the side of the paper in the usual algorithm.)
"About $300, because at one place they got $68, so 5 days would make about $300"
"About $148 because there will be more places." (A column addition of 74 + 74 = 148 is written.)

Add 1 point if there is any reference to "estimating or rounding up/down" or that the amounts collected would not be the same at each of the locations or the next time that day of the week comes around; e.g.:

"About $300, because you can't tell for certain how many will give."
"$400, because some places might do better than others."
"I say $50 for each place or $250 because it could be less."

Top possible score for this section of Part 4 is 4 points.

So the top possible score for the entire Part 4 is 7 points.

Top possible score for the whole item is 13 points.

APPENDIX D

Hands-on problem-solving tasks with solution score coding schemes

Measurement/Body Math: Detective Task (Pretest)

[A corner of the room is set up as follows, to simulate a crime scene (note: measurements are for reference; they are not told to children, although children can choose to measure the objects themselves):

- *A 5-foot-wide frame sits on the floor in front of the wall. Torn paper hangs around the inside edge of the frame, to simulate a picture that's been cut out. There are matching dirty handprints on the left and right edges of the frame where someone held it.*
- *A pair of 10-inch-long footprints face the wall.*
- *A hat (20" around) lies nearby, with one brown hair (represented by yarn) inside]*

Detective task, part 1:

In this puzzle, we're going to pretend that I'm a police officer, and you two are special detectives who are visiting my city to help solve a very mysterious crime. Famous detectives like Sherlock Holmes can tell a lot about crooks from the clues they leave behind. For example, they can use footprints to figure out how tall someone is, how much they weigh, whether they walk with a limp, and so on. Fingerprints can tell them even more, because everybody's fingerprints are different. But in <u>this</u> crime, the crook didn't leave any fingerprints behind, although there are some other kinds of clues you can use. Ready?

Here's what happened: Imagine this is a museum where a priceless painting was stolen. *[While demonstrating:]* The thief came in at night, stood here *[in footprints]*, and took the painting off the wall *[one hand on either side of frame]*, leaving these dirty handprints on the frame. Then, the thief cut out the painting, rolled it up, and ran away. But during the getaway, the thief's hat fell off. It's lying right here, with a blonde hair inside.

I need to catch the thief who took the painting, but I don't know who did it. So I need your help. Your job is to use these clues to figure out as much as you can about what the thief looks like: how tall the thief is, what color hair the thief has, and so on. You might be able to tell a lot about the thief, or maybe just a little. Either way, each thing you figure out will help me narrow down the search, and that'll make it easier for me to catch the crook. So each piece of information is important.

To help you, you can use anything you want from this kit of detective tools – or you don't have to use any of the tools at all.

Okay, now I'm going to give you a little time to figure out the puzzle. You can do whatever you want to help you figure it out, and if you want to use any of this stuff [kit], you can. I'll be over there working if you need me. Otherwise, when you're ready, you can call me and we'll talk about what you think the thief looks like. Any questions? Okay, let's begin.

1. [When kids ready, ask:] What did you figure out about what the thief looks like? Anything else? How do you know the thief looks like that?

Detective task, part 2:

Thanks – that was a great help. But now, I have another problem: While you were working, I got a call on my police radio, saying that a bunch of other museums were also robbed at the same time! The same person couldn't be in all of those places at the same time, so now I have to catch a lot of thieves – and they all left different kinds of footprints and handprints behind. [SHOW 5-6 FOOTPRINTS OF VARIOUS SIZES.]

It's going to take me weeks and weeks to check out all of those crimes. Since you're just visiting my city today, you won't be here to help me with all of them. So now, I need you to teach me how to use these kinds of clues to figure out what people look like, just like you did. Remember, I don't want you to come up with a description of one person this time. Instead, I want you to give me some instructions that I can use for any crime like this, so I can use the same steps every time.

Take a little time to think about it. When you're ready, let me know and we'll talk about what you're thinking. Any questions? Okay, let's begin.

2. [When kids ready, ask:] What are some of the things I can do to figure out what all of the thieves look like? How would that help me figure it out? What steps should I use to figure it out? Anything else?

3. Okay, you said I can figure everything out by [summarize ideas/steps in kids' answer to #2]. Is that the same way you figured out this mystery with the stolen painting? [If no:] What did you do that was different? [If yes:] Was there anything else that you also tried, but it didn't work? [If yes:] What did you try? Did it work? Anything else?

4. [If kids have trouble describing a general procedure in question #2:] Now, let's talk about what you did when you were figuring out this mystery with the painting. I was over there while you were working. Can you tell me what you were doing and what you were thinking about? What did you do first? [Continue with standard probes]

Detective task: Solution score coding scheme

Part 1 (deductions about this particular thief):
Maximum score = 8

- Add 1 point for at least one reasonable, non-mathematical deduction about the thief's appearance (e.g., brown hair, "It's a girl because the hair is long"). (Note: Do not award more than 1 point if child makes more than one such deduction.)
- As child describes the size of the thief's foot, hand, and/or head (i.e., not as child draws additional inferences about other body parts), add EITHER:
    - 0 points for no use of either standard or nonstandard measurement. OR:
    - 1 point if child eyeballs to informally compare the thief's footprint, handprint, or hat size to child's own body (without any more formal use of standard or nonstandard measurement). For example, "The thief's feet are almost the same

size as mine." (Note: Do not award more than 1 point if child does this, e.g., twice for foot and hand.) OR:

  - 2 points if child uses standard or nonstandard measurement <u>inaccurately</u> to document size of footprint, handprint, or hat (<u>without</u> any accurate measurement at other points during the task). For example, child lines up ruler or string inaccurately, or reports measurements that exceed ½" margin of error. (Note: Do not award more than 2 points if child measures more than one clue inaccurately.) OR:
  - 3 points for at least one instance of using standard or nonstandard measurement <u>accurately</u> to document size of footprint, handprint, or hat (within ½" margin of error). For example, child accurately measures footprint with ruler and concludes that the thief's foot is 10" long (within ½" margin of error). Or child uses nonstandard measurement and records result without using numbers, e.g., by cutting or marking string to keep a record of the length of the footprint. (Note: Do not award more than 3 points if child measures more than one clue accurately.)

- In addition, if child uses the above information to draw inferences about the thief's height or size of other body parts (not foot, hand, or head), add EITHER:
  - 0 points for no inferences beyond hand, foot, and head. OR:
  - 1 point for making at least one inference about the thief's body that is based on informally comparing size of clues to own body (<u>without</u> any more formal use of body math). For example, "The thief is either short or a kid because my feet are almost as big as his feet." (Note: Do not award more than 1 point if child does this more than once.) OR:
  - 2 points for at least one <u>inaccurate</u> use of body math to infer size of other body parts or overall height (<u>without</u> any accurate uses of body math at other points during the task). That is, award 2 points if child attempts to use body math, but all attempts are inaccurate. For example, child believes that height is equal to five footprints or twice around the head, or child's conclusion about size exceeds ½" margin of error. (Note: Do not award more than 2 points if child does this more than once.) OR:
  - 3 points for one <u>accurate</u> use of body math to infer size of other body parts or overall height (within ½" margin of error). For example, child uses arm span, foot size, or head size to conclude that the thief is five feet tall (within ½" margin of error). Or child recognizes that the length of the thief's forearm is the same as the thief's foot. Or child draws an accurate inference via nonstandard measurement without using numbers (e.g., cutting or marking a string to equal arm span, and then explaining that the thief is as tall as the string).
  - 4 points for more than one <u>accurate</u> use of body math (within ½" margin of error). For example, child accurately uses body math to infer both overall height and size of forearm.

Part 2 (instructions for solving other, similar crimes):
Maximum score = 6

- Add EITHER:

- 0 points for "don't know" or no response. OR:
- 1 point for non-mathematical response, with no mention of measurement (e.g., "Look hard for clues with a magnifying glass" or "Write down what you figure out," with no mention of measurement). OR:
- 2 points for response that includes at least one idea re: informally comparing size of clues to own body parts, without mentioning either more formal measurement or inferences beyond foot, hand, or head (e.g., "See if the footprints are bigger or smaller than yours"). OR:
- 3 points for using informal comparison (without measurement) to draw reasonable inferences about other body parts (e.g., "If the footprints are bigger than yours, then the thief is probably taller than you"). OR:
- 3 points for at least one idea re: measuring clues, without any mention of inferences beyond foot, hand, or head (e.g., "Use a ruler/string to see how big the footprints are"). OR:
- 4 points for at least one inaccurate or incomplete attempt at general principle that uses body math to draw inferences about other body parts (e.g., "Measure the hat from front to back [instead of circumference], and multiply by 3 to figure out how tall the thief is"), without any accurate attempts. OR:
- 5 points for <u>one</u> accurate general principle that uses body math to draw inferences about other body parts (e.g., "Measure from fingertip to fingertip, and that's the thief's height," "The length of the thief's foot is the same as the length of his forearm"). OR:
- 6 points for <u>more</u> than one accurate general principle that uses body math to draw inferences about other body parts (e.g., "Measure from fingertip to fingertip, and that's the thief's height" <u>and</u> "The length of the thief's foot is the same as the length of his forearm").

Measurement/Body Math: Shaquille O'Neal Task (Posttest)

Shaquille O'Neal task, part 1:

Have you ever gone to a wax museum? That's a museum where they have statues of lots of famous people – actors, singers, athletes, etc. – and the statues are made out of wax.

Let's imagine that I'm a sculptor who makes the statues in a wax museum, and the two of you are visiting me for the day to help me. The next statue I'm supposed to make is Shaquille O'Neal, the basketball player. (SHOW PHOTO) But I have a problem. Usually, when I make a statue of someone, the person comes to the museum so I can measure them and make sure everything on the statue is just right. But Shaq can't come now, because his team has a big basketball game someplace else. All he could send us instead was this photo and one of his sneakers. Here's an outline of his sneaker. (SHOW OUTLINE)

The photo shows what Shaq looks like, but I still need to make sure to make the statue the right size – all the parts of its body need to be exactly the same size as the real Shaq's body.

So I need your help. Your job is to use these clues to figure out as much as you can about Shaq: how tall he is, how big the different parts of his body are, and so on. Then, you'll tell me what you figured out, so that I can make the statue. You might be able to tell a lot about Shaq, or maybe just a little. Either way, each thing you figure out will help me make the statue the right size. So each piece of information is important.

To help you, you can use anything you want from this kit of tools – or you don't have to use any of the tools at all.

Okay, now I'm going to give you a little time to figure out the puzzle. You can do whatever you want to help you figure it out, and if you want to use any of this stuff [*kit*], you can. I'll be over there working if you need me. Otherwise, when you're ready, you can call me and we'll talk about what you figured out about Shaq. Any questions? Okay, let's begin.

1. [When kids ready, ask:] What did you figure out about the statue? How big do you think the different parts of the statue should be? Anything else? How do you know the statue [OR parts of the statue] should be that size [OR those sizes]? [If just give height:] Were you able to figure out how big the different parts of his body should be too?

Shaquille O'Neal task, part 2:

Thanks – that was a great help. But now, I have another problem: Since they heard what Shaq did, lots of other famous people have decided that don't have time to sit around posing either. They just want to send photos of themselves and outlines of <u>their</u> shoes too! [SHOW 5-6 OUTLINES OF SHOES OF VARIOUS SIZES.]

It's going to take me weeks and weeks to make all of those statues. Since you're just visiting the museum for today, you won't be here to help me with all of them. So now, I need you to teach me how to use these kinds of clues to figure out someone's size, just like you did. Remember, I don't want you to come up with a description of one person this time. Instead, I want you to give me some instructions that I can use with any photo and footprint, so I can use the same steps every time I make a new statue.

Take a little time to think about it. When you're ready, let me know and we'll talk about what you're thinking. Any questions? Okay, let's begin.

2. [When kids ready, ask:] What are some of the things I can do to figure out how big to make the parts of each of the statues? How would that help me figure it out? What steps should I use to figure it out? Anything else?

3. Okay, you said I can figure everything out by [summarize ideas/steps in kids' answer to #2]. Is that the same way you figured out how to make your statue of Shaq? [If no:] What did you do that was different? [If yes:] Was there anything else that you also tried, but it didn't work? [If yes:] What did you try? Did it work? Anything else?

4. [If kids have trouble describing a general procedure in question #2:] Now, let's talk about what you did when you were figuring out the Shaq statue. I was over there while you were working. Can you tell me what you were doing and what you were thinking about? What did you do first? [Continue with standard probes]

Shaquille O'Neal task: Solution score coding scheme

Part 1 (making statue of Shaq):
Maximum score = 8

Note: In real life, Shaq's foot is proportionally larger than would be expected for his body. To make the math work correctly in this task, we reduced the size of his footprint to approximately 12.5" long, which yields a total height of 7' 3.5" (close to his actual height, which is either 7' 1" or 7' 2", depending on which source you consult). With that in mind, we took both Shaq's actual measurements and the measurements derived from body math into account in creating the coding scheme below.

- Add 1 point if child uses non-mathematical means to give Shaq's height (7' 1") within one-inch margin of error, or his shoe size (22 EEE) within one size margin of error – e.g., because they read his biography, saw a program about him on TV, or guessed because they knew that some pro basketball players are about 7 feet tall. (Note: Do not award more than 1 point if child gives both Shaq's height and shoe size for these sorts of reasons.)
- As child describes the size of the Shaq's foot (i.e., not as child draws additional inferences about height or other body parts), add EITHER:
    - 0 points for no use of either standard or nonstandard measurement. OR:
    - 1 point if child eyeballs to informally compare Shaq's footprint to child's own body (without any more formal use of standard or nonstandard measurement).

For example, “Shaq’s feet are bigger than mine.” (Note: Do not award more than 1 point if child does this more than once, e.g., for both children in the dyad.) OR:

  - 2 points if child uses standard or nonstandard measurement inaccurately to document size of footprint or photo (without any accurate measurement at other points during the task). For example, child lines up ruler or string inaccurately, or reports measurements that exceed ½” margin of error. (Note: Do not award more than 2 points if child measures more than one clue inaccurately.) OR:
  - 3 points for at least one instance of using standard or nonstandard measurement accurately to document size of footprint, handprint, or hat (within ½” margin of error). For example, child accurately measures footprint with ruler and concludes that the Shaq’s footprint is 12.5” long (within ½” margin of error). Or child uses nonstandard measurement and records result without using numbers, e.g., by cutting or marking string to keep a record of the length of the footprint. (Note: Do not award more than 3 points if child measures more than one clue accurately.)

- In addition, if child uses the above information to draw inferences about Shaq’s height or size of other body parts (not foot), add EITHER:
  - 0 points for no inferences beyond foot. OR:
  - 1 point for making at least one inference about Shaq’s body that is based on informally comparing size of clues to own body (without any more formal use of body math). For example, “Shaq is a lot taller than me because his feet are bigger than my feet.” (Note: Do not award more than 1 point if child does this more than once.) OR:
  - 1 point for making an inaccurate guess about height, etc. without using body math. For example, “Shaq should be 10 feet tall,” “he’s 7 feet tall because I saw it in a book of basketball stats,” “this is 2 inches long, so the statue should be 20 inches long because 2 x 10 = 20 [without any basis for the 10].”
  - 2 points for at least one inaccurate use of body math to infer size of other body parts or overall height (without any accurate uses of body math at other points during the task). That is, award 2 points if child attempts to use body math, but all attempts are inaccurate. For example, child believes that height is equal to five footprints or twice around the head, or child’s conclusion about height exceeds 2” margin of error. (Note: Do not award more than 2 points if child does this more than once.) OR:
  - 3 points for one accurate use of body math to infer size of other body parts or overall height (within 2” margin of error). For example, child uses foot size to conclude that Shaq is 7’ 2” (within 2” margin of error). Or child recognizes that the length of Shaq’s forearm is the same as his foot. Or child draws an accurate inference via nonstandard measurement without using numbers (e.g., cutting or marking a string to equal the length of seven footprints, and then explaining that Shaq is as tall as the string).
  - 4 points for more than one accurate use of body math (within 2” margin of error). For example, child accurately uses body math to infer both overall height and size of forearm.

Part 2 (instructions for making statues of other celebs):
Maximum score = 6

- Add EITHER:
    - 0 points for “don’t know” or no response. OR:
    - 1 point for non-mathematical response, with no mention of measurement (e.g., “Look hard for clues” or “Write down what you figure out,” with no mention of measurement). OR:
    - 2 points for response that includes at least one idea re: informally comparing size of clues to own body parts, without mentioning either more formal measurement or inferences beyond foot (e.g., “See if the footprints are bigger or smaller than yours”). OR:
    - 3 points for using informal comparison (without measurement) to draw reasonable inferences about other body parts (e.g., “If the footprints are bigger than yours, then the person is probably taller than you”). OR:
    - 3 points for at least one idea re: measuring clues, without any mention of body math or inferences beyond the given clues (e.g., “Use a ruler/string to see how big the footprints are”). OR:
    - 4 points for at least one inaccurate or incomplete attempt at general principle that uses body math to draw inferences about other body parts (e.g., “Measure the foot and multiply by 10 to figure out how tall the person is,” “Measure your size [relations] and the other person’s, and if it’s the same for you, it’ll probably be the same for the statue”), without any accurate attempts. OR:
    - 5 points for <u>one</u> accurate general principle that uses body math to draw inferences about other body parts (e.g., “Seven times the footprint is the person’s height,” “The length of the person’s foot is the same as the length of his forearm”). OR:
    - 6 points for <u>more</u> than one accurate general principle that uses body math to draw inferences about other body parts (e.g., “Seven times the footprint is the person’s height” <u>and</u> “The length of the person’s foot is the same as the length of his forearm”).

Organizing Data: Ping Pong Task (Pretest)

[Props for this task consist of a baseball card-like card for each player, and a mock-handwritten note from each player regarding days when he or she is or is not available to play. The notes read:

- *NAME: NATASHA*
  I cannot play ping pong on Tuesday, Wednesday, and Saturday
- *My name is Sally*
  I cannot play on Monday and Thursday.
- *Henry can play on Monday, Wednesday, and Thursday.*
- *I can play on Monday, Tuesday, and Friday.*
  Max]

Ping Pong task, part 1:

Did you ever play ping pong? (IF NO: In ping pong, two people stand on opposite sides of a table with a little net in the middle. The players hit a ball back and forth over the net, and if the other person misses, you get a point.) In this puzzle, we're going to pretend that I'm in charge of organizing a ping-pong tournament at the local recreation center, and the three of you are volunteers who are here to help me. Here’s how the tournament works: There are four players, Natasha, Henry, Sally, and Max (SHOW CARDS), who are going to take turns playing ping pong. In one game, for example, Natasha and Max might play against each other, or maybe Max and Henry. In the tournament, everyone will have a turn to play against everyone else, and then we’ll see who won the most games.

But organizing the tournament is tricky, because not everyone can come to the center every day. Each player has filled out a piece of paper that shows the days when they can or can’t play. (SHOW AND READ THROUGH PAPERS) It’s fine for them to play more than one game on the same day, but they don't want to play any more games than they need for the tournament. And because the recreation center is closed on Sundays, nobody can play on Sunday.

I need your help to set up the tournament. Your job is to figure out a schedule that lets everyone play against everyone else, and also fits the days when each kid can come to the center. Then, you’ll show me what you figured out, so that we can set up the tournament.

Okay, now I’m going to give you a little time to figure it out. You can do whatever you want to help you figure it out, and if you want to use any of this stuff [*kit*], you can. I’ll be over there working if you need me. Otherwise, when you’re ready, you can call me and we’ll talk about what you figured out. Any questions? Okay, let’s begin.

1. [When kids ready, ask:] What did you figure out about the ping pong tournament? How do you think we should set up the schedule? Anything else? How do you know the schedule [OR that day] should be like that? Does this schedule have all of the games we need for the tournament? [IF YES: How do you know?] [IF NO: Why not?]

Ping Pong task, part 2:

Thanks – that was a great help. But now, I have another problem: Later this year, the center is going to have tournaments for other one-on-one games too, like checkers or handball, and I'll have to set up schedules for those tournaments too. But you won't be here all year long to help me set them up.

So now, I need you to teach me how to set up these kinds of schedules, just like you did. Remember, I don't want you to come up with just one schedule now. Instead, I want you to give me some instructions that I can use to set up a schedule for any tournament, so I can use the same steps every time there's a new tournament.

Take a little time to think about it. When you're ready, let me know and we'll talk about what you're thinking. Any questions? Okay, let's begin.

2. [When kids ready, ask:] What are some of the things I can do to figure out how to set up schedules for all of the other tournaments? How would that help me figure it out? What steps should I use to figure it out? Anything else?

3. Okay, you said I can figure everything out by [summarize ideas/steps in kids' answer to #2]. Is that the same way you figured out how to make your schedule for the ping pong tournament? [If no:] What did you do that was different? [If yes:] Was there anything else that you also tried, but it didn't work? [If yes:] What did you try? Did it work? Anything else?

4. [If kids have trouble describing a general procedure in question #2:] Now, let's talk about what you did when you were figuring out the ping pong schedule. I was over there while you were working. Can you tell me what you were doing and what you were thinking about? What did you do first? [Continue with standard probes]

Ping pong task: Solution score

There are two components to a fully correct solution to this task: (1) ensuring that all of the possible matches have been made, and (2) scheduling the matches to fit the days when the players are available. Both of these are taken into account in assigning children a solution score.

Part 1

For making all possible pairs, assign:

- **5 points** if children describe all 6 possible pairs of players with complete explanation, including some systematic way of keeping track of what they are. Create chart of all possible pairs OR use multiplicative or additive ideas, like "Each player needs to play against three other kids..." OR children either compute (4 x 3)/2 (note use of division to avoid double-counting pairs) or add 3 + 2 + 1 to recognize that there are 6 possible pairs.

- **4 points** if children use a systematic strategy (as in a 5-point answer) but make a computational error so that the total does not equal 6 (e.g., make less than 6 pairs, or fail to notice the issue of double-counting and make a total of 12).

- **3 points** if children compile a list of all 6 possibilities, but with no systematic list or process and no adequate justification for having them all (e.g. "I tried all the ways, and that's all there are. I couldn't find any more.")

- **2 points** if children use a trial and error method, randomly assembling an incomplete list of possibilities, or a list with or without duplicates, perhaps drawing examples on paper or placing cards side-by-side.

- **1 point** if children use no multiplicative or additive structure, even informally. Say, for example, that there need to be 4 matches (which is the number of players).

- **0 points** if children give no answer, or children are unable to understand the problem.

For scheduling, assign an additional:

- **4 points** if children assign all 6 pairs of players to days when the relevant players are available, using some systematic means of matching players to available days (e.g., create chart or list of days when each player is available).

- **3 points** if children use a systematic strategy (as in a 5-point answer) but make an error in creating their chart OR only assign 4 or 5 of the possible pairs to days when the players are available.

- **2 points** if children assign all 6 pairs to days when the relevant players are available, using trial and error, guess and check, or other non-systematic strategy (with no chart) in which they randomly assign pairs to days (and perhaps check availability afterward).

- **1 point** if children use a non-systematic method (e.g., trial and error), randomly assigning children to days, and assign at least one pair to a day when the relevant players are available.

- **0 points** if children give no answer, or children are unable to understand the problem.

Note: The days when the players are actually available are:

Mon	Tues	Wed	Thurs	Fri	Sat
Natasha Henry Max	Sally Max	Sally Henry	Natasha Henry	Natasha Sally Max	Sally

Thus, there are several ways to arrange pairs on days when they are available. For example:

Mon	Tues	Wed	Thurs	Fri	Sat
Henry vs. Max	Sally vs. Max	Sally vs. Henry	Natasha vs. Henry	Natasha vs. Sally AND Natasha vs. Max	

or:

Mon	Tues	Wed	Thurs	Fri	Sat
Henry vs. Max AND Natasha vs. Henry AND Natasha vs. Max	Sally vs. Max	Sally vs. Henry		Natasha vs. Sally	

<u>Part 2</u>

As in part 1, points are assigned for a method that ensures making all possible pairs, and for assigning players to days when they are available.

For making all possible pairs, assign:

- **4 points** if children describe a systematic way to keeping track of pairs. Describe way to create chart of all possible pairs OR use multiplicative or additive ideas, like "The first player needs to play against everyone else. The second player already played against the first one, so she needs to play against everyone except the first player..." OR children describe a computational method equivalent to (N x (N-1))/2 (where N = number of players) or (N-1) + (N-2) + … + (N-(N-1)).

- **3 points** if children use a systematic strategy (as in a 5-point answer) but the strategy is flawed due to a conceptual or computational error.

- **2 points** if children describe a non-systematic method (e.g., trial and error) that would succeed in creating some pairs but does not guarantee making all possible pairs.

- **1 point** if children use no multiplicative or additive structure, even informally. Say, for example, to count the number of players).

- **0 points** if children give no answer, or children are unable to understand the problem.

For scheduling, assign an additional:

- **4 points** if children describe a systematic method that would succeed in assigning all pairs of players to days when the relevant players are available (e.g., create chart or list of days when each player is available).

- **3 points** if children describe a systematic strategy (as in a 5-point answer) but their method is flawed and would not succeed in assign all pairs to days when the players are available.

- **2 points** if children describe a non-systematic method (e.g., trial and error) – with no chart -- that would succeed in assigning all pairs to days when the relevant players are available.

- **1 point** if children describe a non-systematic method (e.g., trial and error), randomly assigning children to days, that would succeed in assigning at least one pair to a day when the relevant players are available.

- **0 points** if children give no answer, or children are unable to understand the problem.

Organizing Data: Soccer Task (Posttest)

[Props for this task consist of a baseball card-like card for each team, and a mock-handwritten note regarding days when each team is or is not available to play. The notes read:

- *The Tigers can't play at 2 pm, 5 pm, or 6 pm on Saturday. Tigers rule!*
- *Our team (the #1 Kicks) are ready for Saturday. But we can't play at 1 pm or at 4 pm because two players have dentist appointments.*
- *The only times we can play are at 1 pm, 5 pm, and at 6 pm. Our team is the Comets.*
- *We are ready to play at 2 pm, 4 pm, or 5 pm. See you there!]*

Soccer task, part 1:

Did you ever play soccer? (IF NO: In soccer, two teams play on a field, and they score points by kicking or knocking the ball into each other's goals.) In this puzzle, we're going to pretend that I'm in charge of a soccer league. I need to set up a tournament on Saturday at the local soccer field, and the three of you are volunteers who are here to help me. Here's how the tournament works: There are four teams, the Tigers, the Kicks, the Cheetahs, and the Eagles (SHOW CARDS), who are going to take turns playing against each other. In one game, for example, the Tigers might play against the Eagles, or maybe the Tigers would play against the Kicks. In the tournament, each team will have a turn to play against all of the other teams, and then we'll see which team won the most games.

The soccer field is available for games starting from 1 pm in the afternoon until 6 pm, and each game will take an hour to play. But organizing the tournament is tricky, because not every team can come to the soccer field at the same time of day. The team captains filled out these pieces of paper to show the times when their team can or can't play. (SHOW AND READ THROUGH PAPERS) The teams don't want to play any more games than they need for the tournament.

I need your help to set up the tournament. Your job is to figure out a schedule that lets everyone play against everyone else, and also fits the time of day when each team can come to the field. Then, you'll show me what you figured out, so that we can set up the tournament.

Okay, now I'm going to give you a little time to figure it out. You can do whatever you want to help you figure it out, and if you want to use any of this stuff [*kit*], you can. I'll be over there working if you need me. Otherwise, when you're ready, you can call me and we'll talk about what you figured out. Any questions? Okay, let's begin.

1. [When kids ready, ask:] What did you figure out about the soccer tournament? How do you think we should set up the schedule? Anything else? How do you know the schedule [OR that day] should be like that? Does this schedule have all of the games we need for the tournament? [IF YES: How do you know?] [IF NO: Why not?]

Soccer task, part 2:

Thanks – that was a great help. But now, I have another problem: Later this year, the league is going to have tournaments for other teams from all over the country, and I'll have to set up schedules for those tournaments too. But you won't be here all year long to help me set them up.

So now, I need you to teach me how to set up these kinds of schedules, just like you did. Remember, I don't want you to come up with just one schedule now. Instead, I want you to give me some instructions that I can use to set up a schedule for any tournament, so I can use the same steps every time there's a new tournament.

Take a little time to think about it. When you're ready, let me know and we'll talk about what you're thinking. Any questions? Okay, let's begin.

2. [When kids ready, ask:] What are some of the things I can do to figure out how to set up schedules for all of the other tournaments? How would that help me figure it out? What steps should I use to figure it out? Anything else?

3. Okay, you said I can figure everything out by [summarize ideas/steps in kids' answer to #2]. Is that the same way you figured out how to make your schedule for the soccer tournament? [If no:] What did you do that was different? [If yes:] Was there anything else that you also tried, but it didn't work? [If yes:] What did you try? Did it work? Anything else?

4. [If kids have trouble describing a general procedure in question #2:] Now, let's talk about what you did when you were figuring out the soccer schedule. I was over there while you were working. Can you tell me what you were doing and what you were thinking about? What did you do first? [Continue with standard probes]

Soccer task: Solution score coding scheme

There are two components to a fully correct solution to this task: (1) ensuring that all of the possible matches have been made, and (2) scheduling the matches to fit the times when the players and the field are available. Both of these are taken into account in assigning children a solution score.

Part 1

For making all possible pairs, assign:

- **5 points** if children describe all 6 possible pairs of teams with complete explanation, including some systematic way of keeping track of what they are. Create chart of all possible pairs OR use multiplicative or additive ideas, like "Each team needs to play against three other teams..." OR children either compute (4 x 3)/2 (note use of division to avoid double-counting pairs) or add 3 + 2 + 1 to recognize that there are 6 possible pairs.

- **4 points** if children use a systematic strategy (as in a 5-point answer) but make an error (computational or otherwise) so that the total does not equal 6 (e.g., make less than 6 pairs, or fail to notice the issue of double-counting and make a total of 12).

(Note that drawing a complete list or matrix of pairings or assembling cards to produce a complete list or matrix of pairings constitutes 'a systematic way', whether the result is 6 pairings or not).

- **3 points** if children compile a list of all 6 possibilities, but with no systematic list or process and no adequate justification for having them all (e.g. "I tried all the ways, and that's all there are. I couldn't find any more.")

- **2 points** if children use a trial and error method, randomly assembling an incomplete list of possibilities, or a list with or without duplicates, perhaps drawing examples on paper or placing cards side-by-side – without evidence of a systematic approach for generating pairs.

- **1 point** if children use no multiplicative or additive structure, even informally. Say, for example, that there need to be 4 matches (which is the number of teams).

- **0 points** if children give no answer, or children are unable to understand the problem.

For scheduling, assign an additional:

- **4 points** if children assign all 6 pairs of teams to times when the relevant teams are available, using some systematic means of matching teams to available times (e.g., create chart or list of times when each team is available).

- **3 points** if children use a systematic strategy (as in a 4-point answer) but make an error in creating their chart OR only assign 4 or 5 of the possible pairs to times when the players are available.

- **2 points** if children assign all 6 pairs to times when the relevant players are available, using trial and error, guess and check, or other non-systematic strategy (with no chart) in which they randomly assign pairs to times (and perhaps check availability afterward).

- **1 point** if children use a non-systematic method (e.g., trial and error), randomly assigning teams to times, and assign at least one pair to a time when the relevant teams are available.

- **0 points** if children give no answer, or children are unable to understand the problem.

Note: The times when the teams are actually available are:

1 pm	2 pm	3 pm	4 pm	5 pm	6 pm
Tigers Comets	Kicks Eagles	Tigers Kicks	Tigers Eagles	Kicks Comets Eagles	Kicks Comets

There is only one way to arrange all the pairs at times when they are available:

1 pm	2 pm	3 pm	4 pm	5 pm	6 pm
Tigers v. Comets	Kicks v. Eagles	Tigers v. Kicks	Tigers v. Eagles	Comets v. Eagles	Kicks v. Comets

(Obviously, the order in which teams are listed within a match doesn't matter – e.g., Tigers v. Comets or Comets v. Tigers.)

<u>Part 2</u>

As in part 1, points are assigned for a method that ensures making all possible pairs, and for assigning teams to times when they are available.

For making all possible pairs, assign:

- **4 points** if children describe a systematic way to keeping track of pairs. Describe way to create a chart or list of all possible pairs OR use multiplicative or additive ideas, like "The first team needs to play against everyone else. The second team already played against the first one, so they need to play against everyone except the first team..." OR children describe a computational method equivalent to (N x (N-1))/2 (where N = number of players) or (N-1) + (N-2) + … + (N-(N-1)).

- **3 points** if children use a systematic strategy (as in a 4-point answer) but the strategy is flawed due to a conceptual or computational error.

- **2 points** if children describe a non-systematic method (e.g., trial and error) that would succeed in creating some pairs but does not guarantee making all possible pairs.

- **1 point** if children use no multiplicative or additive structure, even informally. Say, for example, to count the number of players).

- **0 points** if children give no answer, or children are unable to understand the problem.

For scheduling, assign an additional:

- **4 points** if children describe a systematic method that would succeed in assigning all pairs of teams to times when the relevant teams are available (e.g., create chart or list of times when each team is available).

- **3 points** if children describe a systematic strategy (as in a 4-point answer) but their method is flawed and would not succeed in assign all pairs to times when the teams are available.

- **2 points** if children describe a non-systematic method (e.g., trial and error) – with no chart or list -- that would succeed in assigning all pairs to times when the relevant teams are available.

- **1 point** if children describe a non-systematic method (e.g., trial and error), randomly assigning teams to times, that would succeed in assigning at least one pair to a time when the relevant teams are available.

- **0 points** if children give no answer, or children are unable to understand the problem.

APPENDIX E

Hands-on tasks:
Process score coding scheme

Observers identified each of the strategies and heuristics that each triad of children used while working on a given hands-on problem-solving task. Two scores were assigned: a *process score* (equal to the number of heuristics that the triad used while working on the task) and a *unique score* (equal to the number of unique heuristics that the triad used, without duplication). For example, if a triad used a ruler to measure two objects in the course of a task, the heuristic *Measure: ruler* was counted twice in the group's process score, but only once in their unique score. Note that the data presented in the body of this report reflect *unique scores*.

The same coding scheme was used for each of the hands-on tasks. Thus, examples relating to both the body math and ping pong tasks are presented below.

Strategy/heuristic	**Examples**
1: Recall information	Recall of information acquired earlier during same task: "If X worked when I did it before, maybe it would work here too." Recall information from source outside task (e.g., home, school, home, school, *Cyberchase*): "They did something like this on *Cyberchase*…"
2: Gather information	Observes/studies/eyeballs detective clues without moving them around. Studies papers with ping pong players' schedules. Counts number of ping pong players. Asks researcher for clarification of rules or instructions, whether they have been stated/implied or not. (Note: Do not assign points for immediate confirmation. "We need to figure out what the thief looks like, right?" immediately after the introduction of task is confirmation. "Do we need to figure out how tall the thief is?" after measuring the frame is Gather Information.)
3a: Measure: ruler	Uses ruler to measure length.
3b: Measure: count	Counts existing characteristics (e.g., number of pieces in picture frame) <u>as a means of measuring dimensions</u>. (Note that counting for the purpose of finding quantity would be considered Gather Information [2], not 3b.)
3c: Measure: nonstandard manipulatives	Uses paper squares, squares on graph paper, thumb, or other objects in systematic way for measurement. Places own foot in/beside footprint for comparison. Turns frame on side as indication of thief's height, and stands beside it to compare to own height.

Strategy/heuristic	Examples
4: Estimate, approximate	Compares sizes without measuring exactly. Based on size of clues, decides that thief is “bigger than me.” Estimates that it will take “about five or six matches” or “around three days” for all of the ping pong players to play.
5: Calculate	Performs arithmetic calculations, with or without calculator, e.g.: - multiplies to find total number of combinations of ping pong players. - adds or multiplies to compute size of thief via body math.
6a: Manipulate: use objects	Picks up paper footprint to compare to other clues or own body parts. Places “hair” string inside detective notebook as clue. Moves ping pong player cards into pairs to assign matches. Uses objects to represent other objects (e.g., uses paper clips to represent ping pong players).
6b: Manipulate: change objects	Draws or cuts out additional cards for ping pong players in order to make all possible pairs. Cuts string or other object to create measuring tool for measuring body. Draws on pieces (measurements on footprints).

Strategy/heuristic	Examples
7: Trial & error, guess & check	Tries something, recognizes that something is amiss, tries something different, and then checks whether outcome is more satisfactory, e.g.: - measures, then realizes tool wasn't used correctly (e.g., not lined up properly), and tries again. - starts to make pairs of ping pong players, realizes that one of the pairs is double-counted, and either corrects work or starts again. (Note: To be counted as Trial & Error, the new attempt should consist of the same heuristic and be part of the same approach. If a child tries something, then switches to either a different heuristic or different overall approach, then this would be considered Reapproach Problem [11], not Trial & Error.)
8a: Write: List, table, chart	Makes chart or list of ping pong matches. Creates day-by-day calendar page and indicates days when ping pong players are (or are not) available. Writes down list of the thief's characteristics. (Note: To count as a list, the list must have more than one entry. A single entry may be coded under Write: Other [8c].)
8b: Write: Picture, diagram	Draws "wanted" picture of thief without any measurements. Draws calendar page without any indicators of days when ping pong players are available to play. Draws diagram showing the relationship among body parts (e.g., labeled with measurements, or with 7 footprints drawn next to a body – not life-size). (Note: Using pictures for measurement or size comparison also may be coded under Measurement [3a-c] or Manipulate: Change Objects [6b], depending on context.)

Strategy/heuristic	Examples
8c: Write: Other	Writes out equations while performing calculations. (Note: This would also receive a point under Calculate [5].) Writes answer as a means of keeping track. (Note: If this is written as a list – e.g., characteristics of the thief – it would be coded under Write: List, Table, Chart [8a] instead.) Writes notes to self. Writes hash marks to keep track of number of ping pong pairs.
9: Transform problem	Breaks problem into smaller subproblems, e.g.: - focuses on getting all available information from just one detective clue (e.g., footprint) before considering other clues. - first computes total number of possible ping pong pairs, before considering schedule.
10: Look for patterns	Looks for proportional relationships among body parts. Looks for proportional relationships among people (e.g., "You're taller than me, and your hands are bigger too"). Systematically makes all possible pairs of ping pong players.

Strategy/heuristic	**Examples**
11: Reapproach problem	Rejects well-formed approach to the problem in favor of new approach that works better or more efficiently. For example: -begins to measure circumference of wrist with ruler, then switches to flexible, nonstandard measuring tool. (Note: To count as "Reapproach," child must change between methods that fall into two different subcategories of measurement as defined above – Ruler, Count, and Nonstandard. Using, e.g., one flexible tool and switching to the other would be Trial and Error [7], not Reapproach.) -starts to limit description of thief to characteristics that can be measured directly (e.g., length of foot), then realizes can use relations among body parts to figure out more. - starts to make ping pong pairs randomly, or assign them randomly to days, and then switches to more systematic approach. - starts making pairs of ping pong players by moving cards (or simply in their heads), then decides to make chart instead. (Note: Simply switching, e.g., from addition to multiplication as a means of performing the same computation is not sufficient to constitute a "new approach," although it may count as Trial and Error [9], depending on context.)

Strategy/heuristic	Examples
12: Reasonableness	"At first, I thought it was going to be X, but then I realized that couldn't be right…" Checks work to make sure it is correct (e.g., multiplies on paper, then checks with calculator; checks to make sure all pairs of ping pong players have been covered). Decides that proposed solution "can't be right" because, e.g., the thief can't be that tall. Justifies solution as reasonable <u>without</u> prompting from researcher. (Note that justifications in response to researcher requests such as "How do you know?" do not count under Reasonableness, although their content may be coded elsewhere.) (Note: To count as "Reasonableness," child must check/ evaluate <u>a preexisting notion or proposed solution</u>, not simply correct an error in computation along the way. However, a child may get credit for correcting a computational error under "Reasonableness" if noticing the error makes the child realizes that his/her entire solution needs to be adjusted.)
13: Alternative ways to solve	Proposes more than one solution to a problem and considers both correct. Offers more than one way to reach the same solution, e.g.: - explains that the thief's height can be obtained from either the length of the footprint or the space between the handprints. - explains that ping pong players can be matched via either a chart or moving cards around (or two different kinds of charts, if they represent different kinds of solutions).
14: Related problems	Offers way to apply solution to other problems/situations <u>before</u> being asked to do so. (Note: Changes offered <u>after</u> the researcher asks for a general rule do not count as Related Problems, although they may receive points elsewhere.) Suggests <u>new</u>, related problem (e.g., "I wonder what would happen if there were <u>ten</u> people playing ping pong?")

APPENDIX F

Materials in “Toolkit” for Hands-On Tasks

The following tools were available to children during all of the hands-on tasks:

- One magnifying glass
- Two cards marked “Fingerprints” (pretest only)
- One 12” ruler
- One length of string, rolled up (6’ long in the pretest, 8’ long in the posttest)
- One pair of child-friendly scissors
- Two pencils with erasers, plus one red pencil
- Two markers (one dark color and one light color), with medium-fine points
- One small notebook of paper
- One pack of 3” x 3” Post-It™ notes
- One roll of Scotch™ tape
- Ten paper clips (assorted colors)
- Ten square pieces of cardboard, each measuring 1” x 1”
- One calculator

APPENDIX G

Online tracking data:
Coding schemes

RAILROAD REPAIR

Each child who plays the game receives two scores, one reflecting the number of gaps that the child fills correctly (*"Number correct"*) and the other reflecting the sophistication of the strategy that the child used (*"Strategy score"*). Because game play requires more sophisticated thinking as the game progresses, these two scores are not independent. Rather, they are two different ways of tapping into essentially the same construct – i.e., the sophistication of the child's thinking and resulting answers.

Number correct

This score is equal simply to the number of gaps in the tracks that the child fills correctly. Note that this is *not* the same as the number of pieces of track the child places correctly, since a single gap might be filled by two or more pieces (e.g., filling a 1.1 gap by using three pieces of track: .8, .2, and .1).

To figure out the number of gaps a child filled:

- Check the list of gaps in the chart on the next page. That will show you which gaps the child needed to fill on a given screen.
- Compare each gap to the pieces that the child placed successfully. Successful placement is indicated by the word "success" in column K of the spreadsheet. You can see which piece was placed successfully by looking at the corresponding row of column H (e.g., "track4" indicates the piece of track labeled .4).
- If the gap and size of the successful piece of track match (e.g., "track4" with a gap of .4), the child filled the gap. If they do not, check whether several adjacent successes add up to fill the gap (e.g., "track3" and "track1" with a gap of .4).
- Count the number of gaps that the child filled.

Maximum possible score in one round of the game is 22. It can be higher if the child hits the "play again" button to continue beyond one round.

Strategy score

Pilot data show that children's game play typically follows one of three strategies (or a combination of the three):

- *matching* (e.g., filling a .8 gap with a .8 piece of track)
- *additive* (e.g., filling a .8 gap by combining a .7 and a .1 piece of track)
- *reserve pieces* (i.e., thinking strategically about the order in which to use the available pieces, so that the child doesn't run out of the necessary pieces before completing the screen)

The following chart shows the gaps presented on each screen of the game and the least sophisticated strategy that can be used to complete the entire screen:

Screen	Gaps to be filled	Number right	Least sophisticated strategy needed to complete the screen
1	.4	1	Matching
2	.8, .8, .8	4	Additive
3	1.1, 1.1, 1.1	7	Additive
4	1.1, 1.1, 1.1, 1.1	11	Additive
5	.6, .6, .9, 1.2, 1.3	16	Reserve pieces
6	2.0, 2.0	18	Additive
7	1.9, 1.0, 1.4	21	Reserve pieces
8	.8	22	Matching

Note that, although the chart shows the least sophisticated strategy necessary for each screen, more sophisticated strategies can be employed as well. For example, consider the .4 gap on screen 1. A child could use a matching strategy to fill this gap with a .4 piece of track (as most children did). However, he or she could also choose to use an additive strategy and fill the gap with a combination of two pieces of track, e.g., .3 and .1.

The strategy score reflects the *most* sophisticated type of strategy that the child employs over the course of the game. Assign points as follows:

- **3 points** if the child shows evidence of one or more uses of a *reserve pieces* strategy.
- **2 points** if the child shows evidence of one or more uses of an *additive* strategy (without ever using a reserve pieces strategy).
- **1 point** if the child shows evidence of one or more uses of a *matching* strategy (without ever using an additive or reserve pieces strategy).
- **0 points** if the child does not use any of the above strategies, and does not fill any gaps successfully.

Maximum strategy score is 3.

Three sources of information will help you determine the type of strategy (or strategies) the child used: (1) which screens the child completed (since some screens require more sophisticated strategies), (2) the number of pieces a child uses to fill a gap successfully, and (3) any written tips that the child submits via the popup at the end of the game.

We can infer that a child has used a *reserve pieces* strategy if:

- the child successfully completes screen 5 and/or screen 7 (both of which require this strategy), OR
- the child's written tips include a mention of either "saving" pieces (e.g., *"Save the little pieces, so you can make other numbers"*), using pieces of track in a specific order (e.g., *"Use the big pieces first and then the small ones"*), or the order in which to fill the gaps on a screen (e.g., *"Fill the little holes first, so you don't run out of little pieces"*).

We can infer that a child has used an *additive* strategy if:

- the child successfully completes any screen other than screen 1 and screen 8 (since all of the other screens require addition), OR
- the child uses a combination of more than one piece to fill a particular gap, OR
- the child's written tips include a mention of adding pieces (e.g., *"Put together the little pieces to fill the big holes," "Think hard and add them up"*).

We can infer that a child has used a *matching* strategy if:

- the child successfully uses a single piece of track to fill any gap on any screen, OR
- the child's written tips include a mention of matching pieces to gaps (e.g., *"Look carefully and find the right piece to fill the hole"*)

NOTE:

Children often click the "clear all pieces" button (indicated by the word "clear" appearing in column G) when they get stuck. Clicking the "clear all pieces" button does not indicate a particular strategy in itself, but it can signal a point when a child is about to change strategies because the old strategy no longer works.

POUR TO SCORE

Each child who plays the game receives two scores, one reflecting the number of target quantities that the child makes correctly (*"Number correct"*) and the other reflecting the sophistication of the strategy that the child used (*"Strategy score"*). Because game play requires more sophisticated thinking after the child makes the first few quantities, these two scores are not independent. Rather, they are two different ways of tapping into essentially the same construct – i.e., the sophistication of the child's thinking and resulting answers.

Number correct

This score is equal simply to the number of target quantities that the child makes correctly. To find the number of quantities a child made:

- Check the last row of the child's data in column M ("round"). That number will tell you what round the child was on when he or she stopped playing.
- Check the corresponding row of column L ("score") to find the number of correct quantities that the child made in that round.
- If the child played more than one round, use column M to find the last row of each round. Check the corresponding row of column L to find the child's score in each round.
- Add together the number of correct quantities that the child made in each round. This sum is the child's score for *Number correct*.

Maximum score in round 1 is 8. If the child chooses to continue beyond round 1, the maximum score across all three rounds is 24.

Strategy score

Pilot data show that children's game play typically progresses through a sequence of the following strategies (or a combination of the three):

- *matching* (which allows the child to make the first quantity, 8)
- *simple subtractive* (pouring one time from the large container to the small one, which allows the child to make 5 and 3)
- *iterative* (emptying the small container and pouring from the large container to the small one twice, which allows the child to make 2; note that many children get stuck at this point and fail to progress further)
- *advanced* (realizing that, by keeping some liquid in the small container, the child can pour off from the large container to make different quantities; this realization is essential for making the remaining quantities: 7, 4, 1, and 6)

Without the "advanced" strategy, it is largely impossible for children to complete round 1 and move on to subsequent rounds. Once children have acquired this strategy, the same set of strategies can be applied in rounds 2 and 3 as well. Thus, although the *Number correct* score takes all three rounds into account, the *Strategy score* typically can be derived from round 1.

The strategy score reflects the *most* sophisticated type of strategy that the child employs over the course of the game. Assign points as follows:

- **4 points** if the child shows evidence of one or more uses of an *advanced* strategy (reflected in the child successfully making either the quantities 7, 4, 1, or 6 in round 1).
- **3 points** if the child shows evidence of one or more uses of an *iterative* strategy (reflected in the child successfully making the quantity 2 in round 1), without any instances of more sophisticated strategies.
- **2 points** if the child shows evidence of one or more uses of a *simple subtractive* strategy (reflected in the child successfully making either the quantities 5 or 3 in round 1), without any instances of more sophisticated strategies.
- **1 point** if the child shows evidence of one or more uses of a *matching* strategy (reflected in the child successfully making the quantity 8 in round 1), without any instances of more sophisticated strategies.
- **0 points** if the child does not use any of the above strategies, and does not make any quantities successfully.

Maximum strategy score is 4.

The primary source of information for the strategy score is column K ("state"). This column contains a binary variable that shows which quantities the child has made. Before the child makes any quantities, the variable in column K reads 0. When the child makes all of them, it reads:

11111111

In between, 1s indicate the quantities that the child has made successfully, and 0s indicate the quantities that the child has not. The digit indicating 1 quart is at the left of the array and 8 quarts is at the right. For example:

10110011

would mean that the child successfully made the following quantities: 1, 3, 4, 7, and 8.

Note that when the array begins with one or more 0s, Excel does not display the entire array. Instead, it starts with the first digit greater than 0. Thus, for example, the display:

1001

means that the child successfully made the quantities 5 and 8. (In other words, when there are fewer than 8 digits in the array, count down from the far right end of the array to figure out what quantities the 1s represent.)

NOTE:

Clicking the "hint" button causes hash marks to appear on the sides of the two containers, which allows children to shift from estimation to measurement. This is actually a very useful strategy, and most children click the "hint" button at some point. However, because the button is marked "hint" (rather than, e.g., "measure"), it's not clear whether children click the button for this reason, or because they think it will help them in some other way (e.g., by giving them one of the answers). Thus, we should not consider the “hint” button in assigning points for the Strategy score.

SLEUTHS ON THE LOOSE

Each child who plays the game receives two scores, one reflecting the number of correct answers that the child provides (*"Number correct"*) and the other reflecting the sophistication of the strategy that the child used (*"Strategy score"*). As in the rubrics for the other games, these two scores are not independent. Rather, they are two different ways of tapping into essentially the same construct – i.e., the sophistication of the child's thinking and resulting answers.

Number correct

This score is equal simply to the number of times that the child identifies the size of a baby or mama creature correctly. To find the number of correct responses:

- Check the child's final row in column O ("other"). The number in parentheses is the child's number of correct answers. (Note: This is an addition that David made after the pilot test, so there are no numbers in parentheses in column O of the pilot data.)
- If the value in the child's final row of column N ('replay") is greater than 0, the child played the game more than once. In that case, be sure to check column O in both the child's final row *and* his or her final row for the preceding round 3 (which you can identify via column M).
- An alternate method: If the above method fails to work for some reason, you can also count the number of times that the word "correct" appears in column G ("event").

Maximum score across the three rounds is 6. If the child plays more than once, the maximum score could be higher. The correct answers are as follows:

Round	Baby height	Mama height
1	30	120
2	20	100
3	30	60

Strategy score

In our pilot data – and in the absence of exposure to *Cyberchase* material on proportional reasoning and body math -- most children's strategies are fairly rudimentary:

- *copying* (realizing that the first two answers are actually given as an example in the instructions, and copying them; this allows children to produce the first two correct answers, 30 and 120)
- *random guessing* (entering random values, in the hopes that one will be correct)
- *educated guessing* (entering values that are close to the most recent correct answer, in the hopes that one will be correct)

However, last year's pilot testing of hands-on body math tasks showed that some children could apply more sophisticated strategies, particularly if they had seen the *Cyberchase* TV episode about body math. Based on these data, we can extrapolate more sophisticated strategies for the online Sleuths game too:

- *measuring,* without evidence of proportional reasoning (the child uses the on-screen ruler to measure the mama footprint, but does not apply proportional reasoning)
- *inaccurate proportional reasoning* (the child attempts to use measurement and proportional reasoning, but makes a mistake; for example, the child correctly recognizes that the first baby footprint is 10" long and the mama footprint is four times as long, but mistakenly concludes that the mama is 40" tall – note that 40" is actually the length of the mama's footprint, not her height)
- *correct proportional reasoning* (the child uses proportional reasoning correctly to find the height of a baby and/or mama creature)

The strategy score reflects the *most* sophisticated type of strategy that the child employs over the course of the game. Assign points as follows:

- **4 points** if the child shows evidence of one or more uses of a *correct proportional reasoning* strategy.
- **3 points** if the child shows evidence of one or more uses of an *inaccurate proportional reasoning* strategy, without any instances of more sophisticated strategies.
- **2 points** if the child shows evidence of one or more uses of a *measuring* strategy, without any instances of more sophisticated strategies.
- **1 point** if the child shows evidence of one or more uses of an *educated guessing* strategy, without any instances of more sophisticated strategies.
- **0 points** if the child uses a *copying* or *random guessing* strategy, without any instances of more sophisticated strategies, or if the child does not use any strategy at all.

We can infer that a child has used a *correct proportional reasoning* strategy if:

- the child successfully provides more than two correct answers without evidence of random or educated guessing (i.e., without guessing several incorrect answers first), OR
- the child's written tips include a mention or example of correct proportional reasoning, regardless of whether the example is the same as one of the questions in the game (e.g., *"If the baby had a 30 inches foot, the answer would be 120 because 30 x 4 = 120"*).

We can infer that a child has used an *incorrect proportional reasoning* strategy if:

- the child provides incorrect answers that reflect this sort of strategy, as opposed to random or educated guessing (e.g., the child enters a value for the mama's height that is actually the length of the mama's footprint, or the child adds the length of a footprint to the number of footprints instead of multiplying them), OR
- the child's written tips include a mention or example of proportional reasoning that contains an error (e.g., *"A ruler is 12 inches. You move the ruler and count the little feet and multiply by 12"* [instead of multiplying by the length of one footprint]).

We can infer that a child has used a *measuring* strategy if:

- the child extends the ruler to at least the length of the mama footprint (as indicated by the value in column J [“rulerheight”]); in other words:

In round:	**Infer a measuring strategy if the value in column J is at least:**
1	264
2	167
3	231

OR

- the child’s written tips include a mention or example of measurement, without referring to proportional reasoning (e.g., *“Use the ruler,” “See how many footprints make the mama’s footprint”*).
- Note: While thinking, many children play with the ruler randomly, without measuring. Thus, by itself, simply moving or expanding the ruler (as indicated by change in columns H, I, and/or J) is not a sufficient indicator of a measuring strategy. To count as a measuring strategy, the child must expand the ruler to at least one of the values shown in the above table or mention measurement in written tips.

We can infer that a child has used an *educated guessing* strategy if:

- the child enters incorrect values that are close to the most recent correct answer (note that the values can be entered in either a systematic or unsystematic order), OR
- the child’s written tips include a mention of guessing based on a previous correct answer (e.g., *“Guess 10 more than the one before”*).

We can infer that a child has used a *random guessing* strategy if:

- the child enters incorrect values that do not resemble the most recent correct answer, or are not preceded by any correct answers (note that the values can be entered in either a systematic or unsystematic order), OR
- the child’s written tips include a mention of guessing, not based on a previous correct answer (e.g., *“Think hard and guess”*).

We can infer that a child has used a *copying* strategy if:

- the child provides only the first one or two correct answers, without any indication of how those answers were derived, OR
- the child provides one or both of the first two answers shortly after clicking to re-read the instructions (indicated by the word “instructions” in column G), OR
- the child’s written tips include a mention of searching for answers in the instructions, without any written or behavioral indication of a more sophisticated strategy (e.g., *“I looked at the instructions and followed how many inches”*). Note that simply saying, e.g., *“Read the instructions”* (without further elaboration) does not necessarily indicate a copying strategy, since the child could have read the instructions and applied a more sophisticated strategy. Thus, the child’s written tips must be interpreted in light of his or her game play, as indicated above.

APPENDIX H
Paper-and-Pencil Attitude Measures

Interest and Confidence Scale (Pretest)

The following is the pretest version of the interest and confidence scale. A parallel version was administered in the posttest. The posttest version consisted largely of the same items in a new order, with two new non-math items substituted for the existing items, so that children would not feel that they were simply being asked the same questions again and fail to take the measure seriously.

There are lots of different kinds of things that people try to figure out every day. Please circle the answer that shows how <u>you'd</u> feel about trying to figure out the following things:

Figuring out how to take care of a pet.

How interesting would this be for you? (circle your answer)

Very interesting A little interesting So-so A little boring Very boring

How well do you think you could do it? (circle your answer)

Definitely could do it	Maybe could do it	Not Sure	Maybe could <u>not</u> do it	Definitely could <u>not</u> do it

---

Figuring out the speed of something, without using a stopwatch.

How interesting would this be for you? (circle your answer)

Very interesting A little interesting So-so A little boring Very boring

How well do you think you could do it? (circle your answer)

Definitely could do it	Maybe could do it	Not Sure	Maybe could <u>not</u> do it	Definitely could <u>not</u> do it

---

Figuring out whether a game of chance is fair for all of the players.

How interesting would this be for you? (circle your answer)

Very interesting A little interesting So-so A little boring Very boring

How well do you think you could do it? (circle your answer)

Definitely could do it	Maybe could do it	Not Sure	Maybe could <u>not</u> do it	Definitely could <u>not</u> do it

Figuring out a multiplication problem.

How interesting would this be for you? (circle your answer)

Very interesting A little interesting So-so A little boring Very boring

How well do you think you could do it? (circle your answer)

Definitely could do it	Maybe could do it	Not Sure	Maybe could not do it	Definitely could not do it

---

Figuring out how to measure flour for a cake without a using measuring cup.

How interesting would this be for you? (circle your answer)

Very interesting A little interesting So-so A little boring Very boring

How well do you think you could do it? (circle your answer)

Definitely could do it	Maybe could do it	Not Sure	Maybe could not do it	Definitely could not do it

---

Figuring out who'll probably win tomorrow's big game, based on how the teams have played before.

How interesting would this be for you? (circle your answer)

Very interesting A little interesting So-so A little boring Very boring

How well do you think you could do it? (circle your answer)

Definitely could do it	Maybe could do it	Not Sure	Maybe could not do it	Definitely could not do it

---

Figuring out how to measure things without using a ruler.

How interesting would this be for you? (circle your answer)

Very interesting A little interesting So-so A little boring Very boring

How well do you think you could do it? (circle your answer)

Definitely could do it	Maybe could do it	Not Sure	Maybe could not do it	Definitely could not do it

Figuring out the history of your home town.

How interesting would this be for you? (circle your answer)

Very interesting   A little interesting   So-so   A little boring   Very boring

How well do you think you could do it? (circle your answer)

Definitely could do it	Maybe could do it	Not Sure	Maybe could <u>not</u> do it	Definitely could <u>not</u> do it

---

Figuring out ways to keep track of time without using a clock.

How interesting would this be for you? (circle your answer)

Very interesting   A little interesting   So-so   A little boring   Very boring

How well do you think you could do it? (circle your answer)

Definitely could do it	Maybe could do it	Not Sure	Maybe could <u>not</u> do it	Definitely could <u>not</u> do it

---

Figuring out how to set up a fence to make as much space as possible inside.

How interesting would this be for you? (circle your answer)

Very interesting   A little interesting   So-so   A little boring   Very boring

How well do you think you could do it? (circle your answer)

Definitely could do it	Maybe could do it	Not Sure	Maybe could <u>not</u> do it	Definitely could <u>not</u> do it

---

Figuring out the best way to study for a math test.

How interesting would this be for you? (circle your answer)

Very interesting   A little interesting   So-so   A little boring   Very boring

How well do you think you could do it? (circle your answer)

Definitely could do it	Maybe could do it	Not Sure	Maybe could <u>not</u> do it	Definitely could <u>not</u> do it

Figuring out how the sizes of different parts of your body compare to each other.

How interesting would this be for you? (circle your answer)

Very interesting	A little interesting	So-so	A little boring	Very boring

How well do you think you could do it? (circle your answer)

Definitely could do it	Maybe could do it	Not Sure	Maybe could <u>not</u> do it	Definitely could <u>not</u> do it

Motivation Measure

The following measure was designed to assess children's orientations toward motivation to engage and persist in challenging problem-solving tasks. Stories were presented in three contexts: school, computer game, and out-of-school mathematical problem solving. The same stories were presented to boys and girls, with characters given either male (for boys) or female (for girls) names, to avoid gender issues as children selected the characters whom they felt were most like themselves.

Name: ________________________________ I am a (circle one): Girl Boy

My grade (circle one): 3 4 My teacher: ___________________

One day, in math class, the teacher decided to have a contest. She put three hard problems on the board, told the class to split up into teams, and challenged them to try to solve all three problems.

Three friends – Rosie, Shana, and Louise – decided to work together. They figured out the first math problem without too much trouble. But the other problems were harder. They worked on one, then on the other, but after trying each problem a couple of times, they still weren't sure how to figure them out.

Rosie was frustrated. "These problems are too hard!" said Rosie. "Let's stop, and ask the teacher to tell us the answers instead."

Shana disagreed. "We started the problems, so we should try to finish them too," she said. "Let's keep trying."

Louise said, "These problems are hard, but they're kind of interesting. I think we should try again – maybe this time, we'll figure them out."

Which of these three friends sounds the <u>most</u> like you?
(Please circle your choice)

Rosie Shana Louise

Which one sounds the <u>least</u> like you?
(Please circle your choice)

Rosie Shana Louise

If the three of them asked you for your advice, what would you tell them to do?

____

____

____

____

Name: ______________________________   I am a (circle one):   Girl   Boy

My grade (circle one):   3   4   My teacher: _________________

One day, three friends – Melissa, Dorothy, and Ricki – were working together to play a new video game. In the video game, the players were traveling through a strange castle, and they had to figure out clues to find the doors that would let them get from one level of the castle to the next.

The three friends figured out the clues on the first level without too much trouble, and went through the door to the second level. But the problems on the second level were harder. The friends worked on one clue, then on another, but after trying each clue a couple of times, they still weren't sure how to figure them out.

Melissa said, "Whew, these are tough! But we started the game, so we should try to finish it too."

Dorothy was frustrated. "These clues are <u>too</u> hard!" said Dorothy. "My brother Byron plays this game all the time. Let's stop, and ask him to tell us the answers instead."

Ricki said, "Well, the clues are tough, but they're kind of interesting. Let's try again – maybe this time, we'll figure them out."

Which of these three friends sounds the <u>most</u> like you?
(Please circle your choice)

Melissa   Dorothy   Ricki

Which one sounds the <u>least</u> like you?
(Please circle your choice)

Melissa   Dorothy   Ricki

If the three of them asked you for your advice, what would you tell them to do?

____

____

____

____

____

____

Name: ______________________________ I am a (circle one): Girl Boy

My grade (circle one): 3 4 My teacher: _________________

One day, three sisters – Elizabeth, Nina, and Janice – wanted to make an art project as a present for their mother's birthday. For their project, they had to carefully cut out pieces of colored paper and paste them in just the right places on a big piece of cardboard.

The sisters pasted down a lot of pieces of colored paper without too much trouble. But when they tried to add the last couple of pieces, it got harder. They needed to cut the pieces to the exact size that would fill the two empty spaces on the cardboard perfectly. But each time they measured a piece of colored paper and cut it out, it came out either a little too big or a little too small. After trying a few times, they still weren't sure how to figure it out.

Elizabeth said, "Making these pieces is harder than I thought! But it's kind of interesting. Let's try again – maybe this time, we'll figure out how to make them just the right size."

Nina said, "We started this, so we should try to finish it too. Let's keep trying."

But Janice was frustrated. "Making the right sizes is too hard!" said Janice. "Let's stop, and ask someone else to cut them out for us instead."

Which of these three sisters sounds the <u>most</u> like you?
(Please circle your choice)

Elizabeth Nina Janice

Which one sounds the <u>least</u> like you?
(Please circle your choice)

Elizabeth Nina Janice

If the three of them asked you for your advice, what would you tell them to do?

____

____

____

____

____

____

APPENDIX I
Supplementary child and teacher interviews protocols

To help us understand children's experience with and learning from *Cyberchase*, supplemental interviews were conducted with children and teachers after the posttest was completed. These interviews focused on perceptions of children's learning (if any) and reactions to *Cyberchase* itself.

The following are the questions asked in these interviews.

Child Interview

1. Over the past few weeks, you did a lot of stuff with Cyberchase (MENTION ONLY THE MEDIA THAT THE PARTICULAR CHILD USED): You watched some videos, played some hands-on games with your classmates, and also played some games on the Cyberchase Web site. (SHOW SMILEY FACE SCALE) How much did you like doing all of that Cyberchase stuff – was it great, good, ok, not so good, or terrible? Why? (IF NEEDED: Can you point to the face that shows what you thought – was the Cyberchase stuff great, good, ok, not so good, or terrible? Why?)

2. Of all of the Cyberchase things that you watched or played, which ones were your favorites? Why? (IF NEEDED: Which ones did you like best? Why?)

3. Do you think you learned anything from Cyberchase? (IF YES: What kinds of things did you learn? Anything else?) (IF NO: Why not?) (IF CHILD JUST SAYS HE/SHE LEARNED "MATH": What kinds of math did you learn? Anything else?)

4. Over the past few weeks, you did a lot of Cyberchase stuff in school. During that time, did you also do anything outside school, to follow up on ideas or activities from Cyberchase? (IF YES: What did you do? Anything else?)

4a. (IF NEEDED:) For example, did you talk to anyone about things from Cyberchase, or try any activities at home? (IF YES: What did you do? Anything else?)

5. Has Cyberchase helped you in any way – either in school, at home, or anywhere else? (IF YES: How?) (IF NO: Why not?)

6. (SHOW PICTURE OF CYBERCHASE CHARACTERS; LEAVE PICTURE OUT FOR QUESTIONS 6-8) In the Cyberchase videos, you saw these kids solve a lot of problems. How do the kids in Cyberchase figure things out? What sorts of things do they do? How does that help them? Anything else?

7. What about when they try something and it doesn't work, or when something is hard for them to figure out? What do they do then?

8. (POINT TO CHARACTERS) This is Jackie, this is Matt, and this is Inez. Which one of these three kids is most like you? Why?

9. During the weeks when you were using Cyberchase in school, did you also watch the Cyberchase TV show or visit the Cyberchase Web site at home? (IF YES:) What did you do? How often did you do it?? Why?

10. Now that you've finished using Cyberchase in school, would you want to keep watching Cyberchase or visiting the Cyberchase Web site at home? (IF YES:) Why? (IF NO:) Why not?

Teacher Interview

1. How would you rate the educational value of the Cyberchase materials for <u>your students</u>: Great, Good, OK, Not So Good, or Terrible? Why?

2. How would you rate the usefulness of these materials for <u>you</u>: Great, Good, OK, Not So Good, or Terrible? Why?

3. How much fun were the materials for your students: Great, Good, OK, Not So Good, or Terrible? Why?

3a. Which of the Cyberchase materials did the children like most: the videos, the hands-on games, or the games on the Web site? How do you know?

4. Do you think the children learned anything or benefited in any way from using the materials? (IF YES: What/how? Anything else?) (IF NO: Why not?)

5. Within each week, how did you decide when to use the Cyberchase materials?

6. Did you modify or change the Cyberchase activities in any way, or did you basically use them as is? (IF CHANGED:) What kinds of changes did you make? Can you give me an example?

7. Did you have any discussions with your class about some of the math topics in Cyberchase, either beforehand as an introduction, or afterward to follow up? (IF YES:) What kinds of things did you talk about? Can you give me an example?

8. Did you try to connect the Cyberchase activities in any way, or did you just use each one by itself? For example, did you make any connections for the children between the math in the videos and the math in the games, or not? (IF YES:) What kinds of connections did you make? Can you give me an example?

9. How likely would you be to use these Cyberchase materials with other children in the future? Why?

9a. (IF WOULD USE THEM:) Which Cyberchase materials would you use: the video, Web site, hands-on activities, or all of them? Why?

10. Is there anything else you'd like to tell us about either the materials or the children's experience with them?

APPENDIX J
Experimental Statistics

This appendix presents descriptive and inferential statistics for the various quantitative analyses presented in this report. Each statistical table is labeled with the page of the report to which it refers.

Pg. 25 and 27-28: Structural equation modeling: Month-to-month use data

The following table presents the output of path analysis assessing the degree to which children's use of each television series (*Cyberchase* and *SpongeBob Squarepants*) and Web site predicted use in subsequent months. Scores were standardized to control for the children's higher rate of use of *SpongeBob Squarepants* overall.

Each row of the table reports the path estimate for one path. For example, the first row tests the degree to which April viewing of the *Cyberchase* TV series (Apravgcybtv) predicted *Cyberchase* viewing in May (Mayavgcybtv).

Standardized Results for PATH List						
Path			Parameter	Estimate	Standard Error	t Value
Mayavgcybtv	<-	**Apravgcybtv**	Apravgcybtv_Mayavgcybtv	-0.54837	0.21465	-2.55469
Mayavgcybweb	<-	**Apravgcybweb**	Apravgcybweb_Mayavgcybweb	0.58692	0.21841	2.68718
Mayavgsptv	<-	**Apravgsptv**	Apravgsptv_Mayavgsptv	0.81351	0.02619	31.06106
Mayavgspweb	<-	**Apravgspweb**	Apravgspweb_Mayavgspweb	0.55406	0.04580	12.09812
Octavgcybtv	<-	**Mayavgcybtv**	Mayavgcybtv_Octavgcybtv	0.20320	0.06297	3.22721
Octavgcybweb	<-	**Mayavgcybweb**	Mayavgcybweb_Octavgcybweb	0.03922	0.06180	0.63466
Octavgsptv	<-	**Mayavgsptv**	Mayavgsptv_Octavgsptv	0.36305	0.05575	6.51187
Octavgspweb	<-	**Mayavgspweb**	Mayavgspweb_Octavgspweb	0.30766	0.05565	5.52875
Novyavgcybtv	<-	**Octavgcybtv**	Octavgcybtv_Novyavgcybtv	0.88909	0.01891	47.00598
Novyavgcybweb	<-	**Octavgcybweb**	Octavgcybweb_Novyavgcybweb	0.48521	0.05018	9.67020
Novyavgsptv	<-	**Octavgsptv**	Octavgsptv_Novyavgsptv	0.92238	0.01319	69.94865
Novyavgspweb	<-	**Octavgspweb**	Octavgspweb_Novyavgspweb	0.78311	0.02678	29.23926
Decavgcybtv	<-	**Novyavgcybtv**	Novyavgcybtv_Decavgcybtv	0.95097	0.01164	81.67911
Decavgcybweb	<-	**Novyavgcybweb**	Novyavgcybweb_Decavgcybweb	0.86156	0.02073	41.55970
Decavgsptv	<-	**Novyavgsptv**	Novyavgsptv_Decavgsptv	0.94534	0.00918	103.01379
Decavgspweb	<-	**Novyavgspweb**	Novyavgspweb_Decavgspweb	0.82270	0.02353	34.97080
Mayavgcybtv	<-	**Apravgcybweb**	apravgcybweb_Mayavgcybtv	0.64772	0.21324	3.03759
Mayavgcybweb	<-	**Apravgcybtv**	apravgcybtv_Mayavgcybweb	-0.50953	0.21932	-2.32319
Mayavgsptv	<-	**Apravgspweb**	apravgspweb_Mayavgsptv	-0.01222	0.03920	-0.31168
Mayavgspweb	<-	**Apravgsptv**	apravgsptv_Mayavgspweb	0.08706	0.05344	1.62905
Octavgcybtv	<-	**Mayavgcybweb**	mayavgcybweb_Octavgcybtv	0.02336	0.06394	0.36540
Octavgcybweb	<-	**Mayavgcybtv**	mayavgcybtv_Octavgcybweb	0.16803	0.06133	2.73974

Standardized Results for PATH List						
Path			Parameter	Estimate	Standard Error	t Value
Octavgsptv	<-	**Mayavgspweb**	mayavgspweb_Octavgsptv	0.04693	0.05917	0.79317
Octavgspweb	<-	**Mayavgsptv**	mayavgsptv_Octavgspweb	0.06550	0.05756	1.13798
Novyavgcybtv	<-	**Octavgcybweb**	octavgcybweb_Novyavgcybtv	0.03347	0.03005	1.11411
Novyavgcybweb	<-	**Octavgcybtv**	octavgcybtv_Novyavgcybweb	0.15989	0.05431	2.94398
Novyavgsptv	<-	**Octavgspweb**	octavgspweb_Novyavgsptv	0.00765	0.02543	0.30062
Novyavgspweb	<-	**Octavgsptv**	octavgsptv_Novyavgspweb	0.04281	0.03799	1.12671
Decavgcybtv	<-	**Novyavgcybweb**	novyavgcybweb_Decavgcybtv	-0.00352	0.02072	-0.16969
Decavgcybweb	<-	**Novyavgcybtv**	novyavgcybtv_Decavgcybweb	0.02645	0.03036	0.87137
Decavgsptv	<-	**Novyavgspweb**	novyavgspweb_Decavgsptv	0.01693	0.01988	0.85161
Decavgspweb	<-	**Novyavgsptv**	novyavgsptv_Decavgspweb	-0.07108	0.03414	-2.08204

Pg. 27 and 30: Structural equation modeling: Relationships between television and Web use

The following table presents monthly data on the relationship between use of each television series and associated Web site, as well as relationships between use of the two television series, and between use of the two Web sites.

Standardized Results for Covariances Among Errors					
Error of	Error of	Parameter	Estimate	Standard Error	t Value
Apravgcybtv	**Apravgcybweb**	apr_cov1	0.87385	0.03289	26.57140
Apravgcybtv	**Apravgsptv**	apr_cov2	0.73994	0.03222	22.96614
Apravgcybweb	**Apravgsptv**	apr_cov3	0.69634	0.03443	20.22760
Apravgcybtv	**Apravgspweb**	apr_cov4	0.27586	0.05490	5.02457
Apravgcybweb	**Apravgspweb**	apr_cov5	0.34201	0.05238	6.52948
Apravgsptv	**Apravgspweb**	apr_cov6	0.38605	0.05187	7.44309
Mayavgcybtv	**Mayavgcybweb**	May_Cov1	0.36515	0.05224	6.99014
Mayavgcybtv	**Mayavgsptv**	May_Cov2	0.07501	0.03568	2.10245
Mayavgcybweb	**Mayavgsptv**	May_Cov3	-0.02798	0.03599	-0.77742
Mayavgcybtv	**Mayavgspweb**	May_Cov4	0.19456	0.04729	4.11436
Mayavgcybweb	**Mayavgspweb**	May_Cov5	0.12273	0.04856	2.52765
Mayavgsptv	**Mayavgspweb**	May_Cov6	0.09039	0.02946	3.06772
Octavgcybtv	**Octavgcybweb**	Oct_Cov1	0.44328	0.04826	9.18592
Octavgcybtv	**Octavgsptv**	Oct_Cov2	0.21103	0.05365	3.93349

Standardized Results for Covariances Among Errors					
Error of	**Error of**	**Parameter**	**Estimate**	**Standard Error**	**t Value**
Octavgcybweb	**Octavgsptv**	Oct_Cov3	0.12819	0.05551	2.30941
Octavgcybtv	**Octavgspweb**	Oct_Cov4	0.16672	0.05561	2.99813
Octavgcybweb	**Octavgspweb**	Oct_Cov5	0.34266	0.05083	6.74097
Octavgsptv	**Octavgspweb**	Oct_Cov6	0.29960	0.04954	6.04709
Novyavgcybtv	**Novyavgcybweb**	Nov_Cov1	0.11241	0.02204	5.10094
Novyavgcybtv	**Novyavgsptv**	Nov_Cov2	-0.00203	0.00999	-0.20319
Novyavgcybweb	**Novyavgsptv**	Nov_Cov3	-0.0003520	0.01920	-0.01833
Novyavgcybtv	**Novyavgspweb**	Nov_Cov4	0.01660	0.01583	1.04911
Novyavgcybweb	**Novyavgspweb**	Nov_Cov5	0.16188	0.03037	5.32988
Novyavgsptv	**Novyavgspweb**	Nov_Cov6	0.02002	0.01419	1.41044
Decavgcybtv	**Decavgcybweb**	Dec_cov1	0.05508	0.01034	5.32792
Decavgcybtv	**Decavgsptv**	Dec_cov2	0.01704	0.00618	2.75698
Decavgcybweb	**Decavgsptv**	Dec_cov3	0.03044	0.00954	3.18933
Decavgcybtv	**Decavgspweb**	Dec_cov4	0.06951	0.01248	5.56940
Decavgcybweb	**Decavgspweb**	Dec_cov5	0.13414	0.01951	6.87424
Decavgsptv	**Decavgspweb**	Dec_cov6	0.03930	0.01174	3.34671

Pg. 33: General linear modeling: Organizing data task (paper-and-pencil task)

The following tables present statistics comparing the degree of change that the experimental groups demonstrated between the experimental pretest and posttest. Scores were controlled statistically for children's prior use of *Cyberchase* during the naturalistic phase, to ensure that the observed effects were attributable to the experimental treatment and not prior, uncontrolled exposure to *Cyberchase*.

Overall effect of experimental treatment:

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Treatment	4	295	12.13	<.0001
Prior use of Cyberchase	1	295	0.55	0.4590

Significance of within-group change from pretest to posttest (note that "estimate" refers to mean change from pretest to posttest, using scores adjusted for prior use):

Treatment Least Squares Means					
Treatment	Estimate	Standard Error	DF	t Value	Pr > |t|
Control	-1.5126	0.4305	295	-3.51	0.0005
Web only	0.2539	0.3220	295	0.79	0.4311
DVD only	0.7512	0.2446	295	3.07	0.0023
DVD + Web	2.1654	0.4302	295	5.03	<.0001
All	-0.2472	0.3475	295	-0.71	0.4774

Significance of pairwise comparisons in pretest posttest differences (e.g., the first row of the table compares change between the Web Only group and No Exposure control group):

Differences of Treatment Least Squares Means						
Treatment	_Treatment	Estimate	Standard Error	DF	t Value	Pr > |t|
Control	Web only	-1.7664	0.5257	295	-3.36	0.0009
Control	DVD only	-2.2637	0.4836	295	-4.68	<.0001
Control	DVD + Web	-3.6780	0.5617	295	-6.55	<.0001
Control	All	-1.2653	0.5809	295	-2.18	0.0302
Web only	DVD only	-0.4973	0.4015	295	-1.24	0.2164
Web only	DVD + Web	-1.9115	0.5262	295	-3.63	0.0003
Web only	All	0.5011	0.4826	295	1.04	0.3000
DVD only	DVD + Web	-1.4143	0.4822	295	-2.93	0.0036
DVD only	All	0.9984	0.4324	295	2.31	0.0216
DVD + Web	All	2.4126	0.5808	295	4.15	<.0001

Pg. 34: General linear modeling: Measurement task (paper-and-pencil)

The following tables present statistics comparing the degree of change that the experimental groups demonstrated between the experimental pretest and posttest. Scores were controlled statistically for children's prior use of *Cyberchase* during the naturalistic phase, to ensure that the observed effects were attributable to the experimental treatment and not prior, uncontrolled exposure to *Cyberchase*.

Overall effect of experimental treatment:

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Treatment	4	293	2.68	0.0317
Prior use of Cyberchase	1	293	0.05	0.8295

Significance of within-group change from pretest to posttest (note that "estimate" refers to mean change from pretest to posttest, using scores adjusted for prior use):

Treatment Least Squares Means					
Treatment	Estimate	Standard Error	DF	t Value	Pr > |t|
Control	0.3265	0.3270	293	1.00	0.3188
Web only	0.04488	0.2354	293	0.19	0.8490
DVD only	0.5179	0.1711	293	3.03	0.0027
DVD + Web	0.4592	0.3106	293	1.48	0.1404
All	-0.4405	0.2443	293	-1.80	0.0724

Significance of pairwise comparisons in pretest posttest differences (e.g., the first row of the table compares change between the Web Only group and No Exposure control group):

Differences of Treatment Least Squares Means						
Treatment	_Treatment	Estimate	Standard Error	DF	t Value	Pr > |t|
Control	Web only	0.2817	0.3954	293	0.71	0.4768
Control	DVD only	-0.1914	0.3597	293	-0.53	0.5951
Control	DVD + Web	-0.1327	0.4191	293	-0.32	0.7518
Control	All	0.7670	0.4256	293	1.80	0.0726
Web only	DVD only	-0.4730	0.2887	293	-1.64	0.1024
Web only	DVD + Web	-0.4143	0.3824	293	-1.08	0.2794
Web only	All	0.4854	0.3442	293	1.41	0.1596
DVD only	DVD + Web	0.05870	0.3439	293	0.17	0.8646
DVD only	All	0.9584	0.3040	293	3.15	0.0018
DVD + Web	All	0.8997	0.4128	293	2.18	0.0301

Pg. 39: General linear modeling: Organizing data task – process (hands-on)

The following tables present statistics comparing the degree of change that the experimental groups demonstrated between the experimental pretest and posttest. Scores were controlled statistically for children's prior use of *Cyberchase* during the naturalistic phase, to ensure that the observed effects were attributable to the experimental treatment and not prior, uncontrolled exposure to *Cyberchase*.

Overall effect of experimental treatment:

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Treatment	4	18.64	4.21	0.0135

Significance of within-group change from pretest to posttest (note that "estimate" refers to mean change from pretest to posttest, using scores adjusted for prior use):

Treatment Least Squares Means					
Treatment	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Control	0.2329	0.5770	22.61	0.40	0.6903
Web only	0.8440	0.6663	22.66	1.27	0.2181
DVD only	2.8707	0.5505	16.35	5.21	<.0001
DVD + Web	3.5268	0.9460	16.55	3.73	0.0017
All	1.2041	0.5557	16.89	2.17	0.0449

Significance of pairwise comparisons in pretest posttest differences (e.g., the first row of the table compares change between the Web Only group and No Exposure control group):

Differences of Treatment Least Squares Means						
Treatment	**_Treatment**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Control	Web only	-0.6111	0.8814	22.64	-0.69	0.4951
Control	DVD only	-2.6378	0.7975	19.26	-3.31	0.0036
Control	DVD + Web	-3.2939	1.1080	17.93	-2.97	0.0082
Control	All	-0.9713	0.8011	19.55	-1.21	0.2398
Web only	DVD only	-2.0267	0.8643	19.74	-2.34	0.0296
Web only	DVD + Web	-2.6828	1.1571	18.27	-2.32	0.0322
Web only	All	-0.3602	0.8676	19.99	-0.42	0.6825
DVD only	DVD + Web	-0.6561	1.0945	16.5	-0.60	0.5570
DVD only	All	1.6666	0.7822	16.62	2.13	0.0484
DVD + Web	All	2.3226	1.0971	16.63	2.12	0.0497

Pg. 40: General linear modeling: Measurement task – process (hands-on)

The following tables present statistics comparing the degree of change that the experimental groups demonstrated between the experimental pretest and posttest. Scores were controlled statistically for children's prior use of *Cyberchase* during the naturalistic phase, to ensure that the observed effects were attributable to the experimental treatment and not prior, uncontrolled exposure to *Cyberchase*.

Overall effect of experimental treatment:

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Treatment	4	21.99	1.37	0.2767

Significance of within-group change from pretest to posttest (note that "estimate" refers to mean change from pretest to posttest, using scores adjusted for prior use):

Treatment Least Squares Means					
Treatment	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Control	3.5694	1.4603	1.803	2.44	0.1479
Web only	2.2480	1.5292	2.204	1.47	0.2683
DVD only	3.5270	1.4028	1.58	2.51	0.1606
DVD + Web	6.2376	1.8623	4.27	3.35	0.0259
All	3.9123	1.4403	1.696	2.72	0.1344

Significance of pairwise comparisons in pretest posttest differences (e.g., the first row of the table compares change between the Web Only group and No Exposure control group):

Differences of Treatment Least Squares Means						
Treatment	**_Treatment**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Control	Web only	1.3214	1.3328	26.46	0.99	0.3305
Control	DVD only	0.04240	1.1668	21.78	0.04	0.9713
Control	DVD + Web	-2.6682	1.6214	20.3	-1.65	0.1152
Control	All	-0.3429	1.2691	26.58	-0.27	0.7891
Web only	DVD only	-1.2790	1.2610	22.08	-1.01	0.3215
Web only	DVD + Web	-3.9896	1.7729	20.87	-2.25	0.0353

Differences of Treatment Least Squares Means						
Treatment	_Treatment	Estimate	Standard Error	DF	t Value	Pr > |t|
Web only	All	-1.6642	1.2887	24.49	-1.29	0.2086
DVD only	DVD + Web	-2.7106	1.6350	18.55	-1.66	0.1142
DVD only	All	-0.3853	1.1581	19.94	-0.33	0.7428
DVD + Web	All	2.3253	1.7635	21.3	1.32	0.2013

Pgs. 40-42: General linear modeling: Analyses of individual heuristics

Reasonableness:

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Treatment	4	341	7.35	<.0001
time	1	341	25.81	<.0001
Treatment*time	4	341	2.01	0.0922

Differences of Treatment*time Least Squares Means								
Treatment	time	_Treatment	_time	Estimate	Standard Error	DF	t Value	Pr > |t|
All	post	All	pre	1.2722	0.2694	341	4.72	<.0001
All	post	Control	post	0.1503	0.2675	341	0.56	0.5745
All	post	DVD + Web	post	-0.6384	0.2701	341	-2.36	0.0187
All	post	DVD only	post	-0.2236	0.2218	341	-1.01	0.3141
All	post	Web only	post	0.3541	0.3072	341	1.15	0.2498
All	pre	Control	pre	-0.9474	0.2891	341	-3.28	0.0012
All	pre	DVD + Web	pre	-1.3074	0.3275	341	-3.99	<.0001
All	pre	DVD only	pre	-0.8688	0.2556	341	-3.4	0.0008
All	pre	Web only	pre	-0.2611	0.3225	341	-0.81	0.4188
Control	post	Control	pre	0.1745	0.2874	341	0.61	0.5442
Control	post	DVD + Web	post	-0.7887	0.2949	341	-2.67	0.0078
Control	post	DVD only	post	-0.3739	0.2515	341	-1.49	0.1379
Control	post	Web only	post	0.2038	0.3292	341	0.62	0.5362
Control	pre	DVD + Web	pre	-0.36	0.3214	341	-1.12	0.2634
Control	pre	DVD only	pre	0.0786	0.2477	341	0.32	0.7511
Control	pre	Web only	pre	0.6863	0.3163	341	2.17	0.0307
DVD + Web	post	DVD + Web	pre	0.6032	0.3281	341	1.84	0.0669
DVD + Web	post	DVD only	post	0.4148	0.2542	341	1.63	0.1037
DVD + Web	post	Web only	post	0.9925	0.3313	341	3	0.0029
DVD + Web	pre	Web only	pre	1.0463	0.3517	341	2.97	0.0031
DVD only	post	Web only	post	0.5777	0.2933	341	1.97	0.0497

DVD only	pre	Web only	pre	0.6077	0.286	341	2.13	0.0343
Web only	post	Web only	pre	0.657	0.3547	341	1.85	0.0649

Recall *Cyberchase*:

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Treatment	4	339	3.04	0.0176
time	1	339	10.95	0.001
Treatment*time	4	339	2.12	0.0775

Differences of Treatment*time Least Squares Means								
Treatment	**time**	**_Treatment**	**_time**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
All	post	All	pre	0.7878	0.3593	339	2.19	0.029
All	post	Control	post	2.1987	0.9139	339	2.41	0.0167
All	post	DVD + Web	post	-0.4397	0.4376	339	-1	0.3157
All	post	DVD only	post	-0.3382	0.3318	339	-1.02	0.3088
All	post	Web only	post	0.2882	0.457	339	0.63	0.5287
All	pre	Control	pre	0.7525	0.6192	339	1.22	0.2251
All	pre	DVD + Web	pre	2.0352	1.6202	339	1.26	0.2099
All	pre	DVD only	pre	0.4031	0.4149	339	0.97	0.3319
All	pre	Web only	pre	1.996	0.9401	339	2.12	0.0345
Control	post	Control	pre	-0.6584	1.0438	339	-0.63	0.5286
Control	post	DVD + Web	post	-2.6385	0.9453	339	-2.79	0.0055
Control	post	DVD only	post	-2.5369	0.9012	339	-2.82	0.0052
Control	post	Web only	post	-1.9105	0.9544	339	-2	0.0461
Control	pre	DVD + Web	pre	1.2827	1.6981	339	0.76	0.4506
Control	pre	DVD only	pre	-0.3494	0.6561	339	-0.53	0.5947
Control	pre	Web only	pre	1.2435	1.0688	339	1.16	0.2455
DVD + Web	post	DVD only	post	0.1016	0.4102	339	0.25	0.8046
DVD + Web	post	Web only	post	0.7279	0.5168	339	1.41	0.1599
DVD + Web	pre	DVD only	pre	-1.632	1.6347	339	-1	0.3188
DVD + Web	pre	Web only	pre	-0.03921	1.8395	339	-0.02	0.983
DVD only	post	DVD only	pre	1.5291	0.3912	339	3.91	0.0001
DVD only	post	Web only	post	0.6264	0.4309	339	1.45	0.1469
DVD only	pre	Web only	pre	1.5928	0.9648	339	1.65	0.0997
Web only	post	Web only	pre	2.4956	0.9816	339	2.54	0.0115

Manipulate: Change objects:

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Treatment	4	341	0.42	0.7969
time	1	341	1.68	0.1956
Treatment*time	4	341	5.3	0.0004

Differences of Treatment*time Least Squares Means								
Treatment	**time**	**_Treatment**	**_time**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
All	post	All	pre	1.0485	0.2982	341	3.52	0.0005
All	post	Control	post	1.0059	0.4198	341	2.4	0.0171
All	post	DVD + Web	post	1.1658	0.629	341	1.85	0.0647
All	post	DVD only	post	0.7419	0.3395	341	2.19	0.0295
All	post	Web only	post	0.7827	0.4224	341	1.85	0.0647
All	pre	Control	pre	-0.8073	0.3141	341	-2.57	0.0106
All	pre	DVD + Web	pre	-0.7402	0.416	341	-1.78	0.0761
All	pre	DVD only	pre	-0.9351	0.2642	341	-3.54	0.0005
All	pre	Web only	pre	-0.3888	0.3228	341	-1.2	0.2293
Control	post	Control	pre	-0.7647	0.4312	341	-1.77	0.0771
Control	post	DVD + Web	post	0.1598	0.6977	341	0.23	0.8189
Control	post	DVD only	post	-0.264	0.4543	341	-0.58	0.5616
Control	post	Web only	post	-0.2232	0.5192	341	-0.43	0.6676
Control	pre	DVD + Web	pre	0.06714	0.423	341	0.16	0.874
Control	pre	DVD only	pre	-0.1278	0.2751	341	-0.46	0.6424
Control	pre	Web only	pre	0.4185	0.3318	341	1.26	0.208
DVD + Web	post	DVD + Web	pre	-0.8575	0.6927	341	-1.24	0.2166
DVD + Web	post	DVD only	post	-0.4238	0.6526	341	-0.65	0.5165
DVD + Web	post	Web only	post	-0.383	0.6993	341	-0.55	0.5843
DVD + Web	pre	DVD only	pre	-0.195	0.3874	341	-0.5	0.6151
DVD + Web	pre	Web only	pre	0.3514	0.4295	341	0.82	0.4139
DVD only	post	DVD only	pre	-0.6286	0.31	341	-2.03	0.0434
DVD only	post	Web only	post	0.04081	0.4567	341	0.09	0.9289
DVD only	pre	Web only	pre	0.5464	0.285	341	1.92	0.0561
Web only	post	Web only	pre	-0.123	0.4401	341	-0.28	0.78

Manipulate: Use objects:

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Treatment	4	341	0.83	0.5065
time	1	341	0.15	0.6965
Treatment*time	4	341	2.79	0.0266

Differences of Treatment*time Least Squares Means								
Treatment	time	_Treatment	_time	Estimate	Standard Error	DF	t Value	Pr > |t|
All	post	All	pre	-0.3747	0.1908	341	-1.96	0.0504
All	post	Control	post	-0.2913	0.236	341	-1.23	0.2179
All	post	DVD + Web	post	-0.2281	0.3076	341	-0.74	0.4589
All	post	DVD only	post	-0.1585	0.222	341	-0.71	0.4758
All	post	Web only	post	-0.2411	0.2532	341	-0.95	0.3416
All	pre	Control	pre	0.5754	0.2036	341	2.83	0.005
All	pre	DVD + Web	pre	0.3365	0.2646	341	1.27	0.2044
All	pre	DVD only	pre	-0.07522	0.1359	341	-0.55	0.5804
All	pre	Web only	pre	0.3365	0.1708	341	1.97	0.0497
Control	post	Control	pre	0.492	0.2465	341	2	0.0467
Control	post	DVD + Web	post	0.06322	0.3082	341	0.21	0.8376
Control	post	DVD only	post	0.1328	0.2229	341	0.6	0.5515
Control	post	Web only	post	0.05016	0.254	341	0.2	0.8435
Control	pre	DVD + Web	pre	-0.2389	0.3065	341	-0.78	0.4363
Control	pre	DVD only	pre	-0.6506	0.2059	341	-3.16	0.0017
Control	pre	Web only	pre	-0.2389	0.2305	341	-1.04	0.3007
DVD + Web	post	DVD + Web	pre	0.1899	0.3581	341	0.53	0.5963
DVD + Web	post	DVD only	post	0.06962	0.2976	341	0.23	0.8152
DVD + Web	post	Web only	post	-0.01306	0.3216	341	-0.04	0.9676
DVD + Web	pre	DVD only	pre	-0.4117	0.2664	341	-1.55	0.1232
DVD + Web	pre	Web only	pre	-0.00001	0.2858	341	0	1
DVD only	post	DVD only	pre	-0.2914	0.177	341	-1.65	0.1006
DVD only	post	Web only	post	-0.08268	0.241	341	-0.34	0.7317
DVD only	pre	Web only	pre	0.4117	0.1736	341	2.37	0.0183
Web only	post	Web only	pre	0.2029	0.2385	341	0.85	0.3954

Gather information:

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Treatment	4	341	3.64	0.0064
time	1	341	1.57	0.2108
Treatment*time	4	341	6.62	<.0001

Differences of Treatment*time Least Squares Means								
Treatment	**time**	**_Treatment**	**_time**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
All	post	All	pre	-0.4596	0.1877	341	-2.45	0.0149
All	post	Control	post	-0.5991	0.2215	341	-2.7	0.0072
All	post	DVD + Web	post	-0.8165	0.2647	341	-3.08	0.0022
All	post	DVD only	post	-0.4576	0.2098	341	-2.18	0.0299
All	post	Web only	post	-0.1897	0.2539	341	-0.75	0.4554
All	pre	DVD + Web	pre	0.6063	0.2796	341	2.17	0.0308
All	pre	DVD only	pre	-0.3389	0.1247	341	-2.72	0.0069
All	pre	Web only	pre	0.1437	0.1555	341	0.92	0.3561
Control	post	Control	pre	0.6216	0.2227	341	2.79	0.0055
Control	post	DVD + Web	post	-0.2174	0.2546	341	-0.85	0.3937
Control	post	DVD only	post	0.1415	0.197	341	0.72	0.4731
Control	post	Web only	post	0.4094	0.2434	341	1.68	0.0935
Control	pre	DVD + Web	pre	0.1243	0.3126	341	0.4	0.6912
Control	pre	DVD only	pre	-0.821	0.1874	341	-4.38	<.0001
Control	pre	Web only	pre	-0.3383	0.2092	341	-1.62	0.1067
DVD + Web	post	DVD only	post	0.3589	0.2445	341	1.47	0.143
DVD + Web	post	Web only	post	0.6268	0.2832	341	2.21	0.0275
DVD + Web	pre	DVD only	pre	-0.9453	0.2785	341	-3.39	0.0008
DVD + Web	pre	Web only	pre	-0.4626	0.2936	341	-1.58	0.116
DVD only	post	Web only	post	0.2679	0.2328	341	1.15	0.2506
DVD only	pre	Web only	pre	0.4827	0.1535	341	3.14	0.0018
Web only	post	Web only	pre	-0.1262	0.2311	341	-0.55	0.5855

Write: List

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Treatment	4	341	0.66	0.6219
time	1	341	8.23	0.0044
Treatment*time	4	341	0.66	0.6206

Differences of Treatment*time Least Squares Means								
Treatment	time	_Treatment	_time	Estimate	Standard Error	DF	t Value	Pr > |t|
All	post	All	pre	0.4145	0.245	341	1.69	0.0915
All	post	Control	post	-0.1451	0.2811	341	-0.52	0.6061
All	post	DVD + Web	post	-0.03175	0.377	341	-0.08	0.9329
All	post	DVD only	post	0.1178	0.2714	341	0.43	0.6646
All	post	Web only	post	-0.3682	0.2839	341	-1.3	0.1955
All	pre	Control	pre	-0.116	0.2669	341	-0.43	0.664
All	pre	DVD + Web	pre	0.1133	0.3987	341	0.28	0.7764
All	pre	DVD only	pre	-0.2386	0.2111	341	-1.13	0.259
All	pre	Web only	pre	-0.1743	0.2405	341	-0.72	0.469
Control	post	Control	pre	0.4436	0.3004	341	1.48	0.1407
Control	post	DVD + Web	post	0.1133	0.3833	341	0.3	0.7677
Control	post	DVD only	post	0.2629	0.2802	341	0.94	0.3488
Control	post	Web only	post	-0.2232	0.2922	341	-0.76	0.4456
Control	pre	DVD + Web	pre	0.2294	0.4294	341	0.53	0.5936
Control	pre	DVD only	pre	-0.1226	0.2645	341	-0.46	0.6433
Control	pre	Web only	pre	-0.0583	0.2886	341	-0.2	0.84
DVD + Web	post	DVD only	post	0.1495	0.3763	341	0.4	0.6913
DVD + Web	post	Web only	post	-0.3365	0.3853	341	-0.87	0.3832
DVD + Web	pre	DVD only	pre	-0.352	0.3972	341	-0.89	0.3761
DVD + Web	pre	Web only	pre	-0.2877	0.4136	341	-0.7	0.4872
DVD only	post	Web only	post	-0.486	0.2829	341	-1.72	0.0867
DVD only	pre	Web only	pre	0.0643	0.2379	341	0.27	0.7871
Web only	post	Web only	pre	0.6084	0.28	341	2.17	0.0305

Estimate:

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Treatment	4	341	0.17	0.9526
time	1	341	1.81	0.18
Treatment*time	4	341	2	0.0942

Differences of Treatment*time Least Squares Means								
Treatment	time	_Treatment	_time	Estimate	Standard Error	DF	t Value	Pr > |t|
All	post	All	pre	-0.2469	0.2037	341	-1.21	0.2264
All	post	Control	post	-0.08178	0.2596	341	-0.32	0.7529
All	post	DVD + Web	post	-0.2837	0.322	341	-0.88	0.3788

All	post	DVD only	post	-0.2125	0.2322	341	-0.91	0.3609
All	post	Web only	post	-0.305	0.2649	341	-1.15	0.2505
All	pre	Control	pre	0.06817	0.1948	341	0.35	0.7266
All	pre	DVD + Web	pre	0.2531	0.2847	341	0.89	0.3746
All	pre	DVD only	pre	0.2913	0.1635	341	1.78	0.0757
All	pre	Web only	pre	0.5408	0.1997	341	2.71	0.0071
Control	post	Control	pre	-0.09692	0.2526	341	-0.38	0.7015
Control	post	DVD + Web	post	-0.2019	0.3317	341	-0.61	0.543
Control	post	DVD only	post	-0.1307	0.2456	341	-0.53	0.595
Control	post	Web only	post	-0.2232	0.2767	341	-0.81	0.4205
Control	pre	DVD + Web	pre	0.1849	0.3115	341	0.59	0.5531
Control	pre	DVD only	pre	0.2231	0.2067	341	1.08	0.281
Control	pre	Web only	pre	0.4726	0.2364	341	2	0.0463
DVD + Web	post	DVD + Web	pre	0.2899	0.3784	341	0.77	0.4441
DVD + Web	post	DVD only	post	0.07126	0.3108	341	0.23	0.8188
DVD + Web	post	Web only	post	-0.02123	0.3359	341	-0.06	0.9496
DVD + Web	pre	DVD only	pre	0.03822	0.2929	341	0.13	0.8963
DVD + Web	pre	Web only	pre	0.2877	0.3146	341	0.91	0.3611
DVD only	post	DVD only	pre	0.2569	0.1979	341	1.3	0.1952
DVD only	post	Web only	post	-0.09249	0.2512	341	-0.37	0.713
DVD only	pre	Web only	pre	0.2495	0.2113	341	1.18	0.2386
Web only	post	Web only	pre	0.5989	0.2619	341	2.29	0.0228

Look for patterns:

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Treatment	4	341	5.87	0.0001
time	1	341	26.68	<.0001
Treatment*time	4	341	2.2	0.0688

Differences of Treatment*time Least Squares Means								
Treatment	**time**	**_Treatment**	**_time**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
All	post	All	pre	0.7717	0.2854	341	2.7	0.0072
All	post	Control	post	-0.7824	0.261	341	-3	0.0029
All	post	DVD + Web	post	-0.9343	0.3047	341	-3.07	0.0023
All	post	DVD only	post	-0.7817	0.2457	341	-3.18	0.0016
All	post	Web only	post	0.5108	0.3939	341	1.3	0.1956
All	pre	Control	pre	-0.5381	0.304	341	-1.77	0.0776
All	pre	DVD + Web	pre	-0.6176	0.3899	341	-1.58	0.1142
All	pre	DVD only	pre	-0.484	0.259	341	-1.87	0.0625
All	pre	Web only	pre	-0.4572	0.2881	341	-1.59	0.1135

Control	post	Control	pre	1.016	0.2813	341	3.61	0.0003
Control	post	DVD + Web	post	-0.1519	0.2754	341	-0.55	0.5815
Control	post	DVD only	post	0.000678	0.2082	341	0	0.9974
Control	post	Web only	post	1.2932	0.3717	341	3.48	0.0006
Control	pre	DVD + Web	pre	-0.07946	0.4084	341	-0.19	0.8458
Control	pre	DVD only	pre	0.05407	0.286	341	0.19	0.8502
Control	pre	Web only	pre	0.08088	0.3126	341	0.26	0.796
DVD + Web	post	DVD only	post	0.1526	0.2609	341	0.59	0.5589
DVD + Web	post	Web only	post	1.4451	0.4036	341	3.58	0.0004
DVD + Web	pre	DVD only	pre	0.1335	0.3761	341	0.36	0.7228
DVD + Web	pre	Web only	pre	0.1603	0.3967	341	0.4	0.6863
DVD only	post	DVD only	pre	1.0694	0.2145	341	4.99	<.0001
DVD only	post	Web only	post	1.2925	0.361	341	3.58	0.0004
DVD only	pre	Web only	pre	0.02681	0.2691	341	0.1	0.9207
Web only	post	Web only	pre	-0.1963	0.3959	341	-0.5	0.6203

Pg. 42-43: General linear modeling: Organizing data task – solution (hands-on)

The following tables present statistics comparing the degree of change that the experimental groups demonstrated between the experimental pretest and posttest. Scores were controlled statistically for children's prior use of *Cyberchase* during the naturalistic phase, to ensure that the observed effects were attributable to the experimental treatment and not prior, uncontrolled exposure to *Cyberchase*.

Overall effect of experimental treatment:

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Treatment	4	19.7	1.27	0.3163

Significance of within-group change from pretest to posttest (note that "estimate" refers to mean change from pretest to posttest, using scores adjusted for prior use):

Treatment Least Squares Means					
Treatment	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Control	1.0743	1.1165	6.349	0.96	0.3711
Web only	2.6422	1.1995	10.66	2.20	0.0506
DVD only	4.1259	0.9803	5.575	4.21	0.0066

Treatment Least Squares Means					
Treatment	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
DVD + Web	3.0537	1.8984	11.74	1.61	0.1343
All	2.3977	1.0995	4.846	2.18	0.0828

Significance of pairwise comparisons in pretest posttest differences (e.g., the first row of the table compares change between the Web Only group and No Exposure control group):

Differences of Treatment Least Squares Means						
Treatment	**_Treatment**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Control	Web only	-1.5679	1.6230	25.74	-0.97	0.3430
Control	DVD only	-3.0516	1.4048	21.18	-2.17	0.0413
Control	DVD + Web	-1.9794	1.9718	19.85	-1.00	0.3275
Control	All	-1.3234	1.6368	20.85	-0.81	0.4279
Web only	DVD only	-1.4837	1.5011	21.55	-0.99	0.3339
Web only	DVD + Web	-0.4115	2.2706	19.92	-0.18	0.8580
Web only	All	0.2445	1.5288	22.28	0.16	0.8744
DVD only	DVD + Web	1.0722	2.0593	18.08	0.52	0.6089
DVD only	All	1.7282	1.4410	18.39	1.20	0.2456
DVD + Web	All	0.6560	2.3633	15.95	0.28	0.7849

Pg. 43: General linear modeling: Measurement task – solution (hands-on)

The following tables present statistics comparing the degree of change that the experimental groups demonstrated between the experimental pretest and posttest. Scores were controlled statistically for children’s prior use of *Cyberchase* during the naturalistic phase, to ensure that the observed effects were attributable to the experimental treatment and not prior, uncontrolled exposure to *Cyberchase*.

Overall effect of experimental treatment:

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Treatment	4	18.63	1.71	0.1891

Significance of within-group change from pretest to posttest (note that “estimate” refers to mean change from pretest to posttest, using scores adjusted for prior use):

Treatment Least Squares Means					
Treatment	Estimate	Standard Error	DF	t Value	Pr > |t|
Control	1.0663	0.9995	2.162	1.07	0.3905
Web only	2.2961	1.0599	2.803	2.17	0.1252
DVD only	2.8228	0.9643	1.955	2.93	0.1022
DVD + Web	3.9163	1.3762	6.09	2.85	0.0289
All	1.7522	0.9925	2.081	1.77	0.2146

Significance of pairwise comparisons in pretest posttest differences (e.g., the first row of the table compares change between the Web Only group and No Exposure control group):

Differences of Treatment Least Squares Means						
Treatment	_Treatment	Estimate	Standard Error	DF	t Value	Pr > |t|
Control	Web only	-1.2298	1.0304	22.27	-1.19	0.2452
Control	DVD only	-1.7564	0.9163	18.63	-1.92	0.0707
Control	DVD + Web	-2.8500	1.2886	17.07	-2.21	0.0409
Control	All	-0.6859	0.9844	21.71	-0.70	0.4933
Web only	DVD only	-0.5266	0.9868	18.88	-0.53	0.5998
Web only	DVD + Web	-1.6201	1.4106	17.91	-1.15	0.2659
Web only	All	0.5439	1.0003	20.1	0.54	0.5926
DVD only	DVD + Web	-1.0935	1.3124	16.13	-0.83	0.4169
DVD only	All	1.0706	0.9169	16.99	1.17	0.2591
DVD + Web	All	2.1641	1.4047	17.9	1.54	0.1409

Pg. 62: Attitudes Toward Mathematics

The following tables present statistics comparing the degree of change that the experimental groups demonstrated between the experimental pretest and posttest. Scores were controlled statistically for children's prior use of *Cyberchase* during the naturalistic phase, to ensure that the observed effects were attributable to the experimental treatment and not prior, uncontrolled exposure to *Cyberchase*.

Interest - *Cyberchase* math:

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
State	1	266	0.91	0.3401
Treatment	4	266	0.70	0.5900

Treatment Least Squares Means							
Treatment	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**	**Mean**	**Standard Error Mean**
Control	-0.4318	0.3674	266	-1.18	0.2409	-0.4318	0.3674
Web only	-0.01693	0.1773	266	-0.10	0.9240	-0.01693	0.1773
DVD only	0.1153	0.1204	266	0.96	0.3392	0.1153	0.1204
DVD + Web	0.02104	0.2160	266	0.10	0.9225	0.02104	0.2160
All	0.1521	0.1433	266	1.06	0.2896	0.1521	0.1433

Interest – non-*Cyberchase* math:

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
State	1	191	0.37	0.5462
Treatment	3	191	0.54	0.6544

Treatment Least Squares Means							
Treatment	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**	**Mean**	**Standard Error Mean**
Control	-0.1599	0.4102	191	-0.39	0.6972	-0.1599	0.4102
Web only	-0.03547	0.2279	191	-0.16	0.8765	-0.03547	0.2279
DVD only	0.02840	0.1856	191	0.15	0.8786	0.02840	0.1856
All	0.2191	0.1686	191	1.30	0.1952	0.2191	0.1686

Differences of Treatment Least Squares Means						
Treatment	_Treatment	Estimate	Standard Error	DF	t Value	Pr > |t|
Control	Web only	-0.1244	0.4423	191	-0.28	0.7788
Control	DVD only	-0.1883	0.4321	191	-0.44	0.6636
Control	All	-0.3790	0.4211	191	-0.90	0.3692
Web only	DVD only	-0.06388	0.2722	191	-0.23	0.8147
Web only	All	-0.2546	0.2559	191	-1.00	0.3210
DVD only	All	-0.1907	0.2306	191	-0.83	0.4093

Interest – school math:

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
State	1	267	0.75	0.3878
Treatment	4	267	6.91	<.0001

Treatment Least Squares Means							
Treatment	Estimate	Standard Error	DF	t Value	Pr > |t|	Mean	Standard Error Mean
Control	-2.5318	0.5826	267	-4.35	<.0001	-2.5318	0.5826
Web only	0.08545	0.2789	267	0.31	0.7596	0.08545	0.2789
DVD only	0.2625	0.1899	267	1.38	0.1680	0.2625	0.1899
DVD + Web	-0.7940	0.3431	267	-2.31	0.0214	-0.7940	0.3431
All	-0.2342	0.2314	267	-1.01	0.3123	-0.2342	0.2314

Differences of Treatment Least Squares Means						
Treatment	_Treatment	Estimate	Standard Error	DF	t Value	Pr > |t|
Control	Web only	-2.6172	0.6286	267	-4.16	<.0001
Control	DVD only	-2.7943	0.6161	267	-4.54	<.0001
Control	DVD + Web	-1.7378	0.7053	267	-2.46	0.0144
Control	All	-2.2976	0.6050	267	-3.80	0.0002
Web only	DVD only	-0.1771	0.3408	267	-0.52	0.6037
Web only	DVD + Web	0.8794	0.4666	267	1.88	0.0605

Differences of Treatment Least Squares Means						
Treatment	_Treatment	Estimate	Standard Error	DF	t Value	Pr > |t|
Web only	All	0.3197	0.3412	267	0.94	0.3497
DVD only	DVD + Web	1.0565	0.3869	267	2.73	0.0067
DVD only	All	0.4967	0.3039	267	1.63	0.1033
DVD + Web	All	-0.5597	0.4452	267	-1.26	0.2098

Interest: Non-math

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
State	1	272	1.31	0.2529
Treatment	4	272	0.45	0.7742

Treatment Least Squares Means							
Treatment	Estimate	Standard Error	DF	t Value	Pr > |t|	Mean	Standard Error Mean
Control	-0.7981	0.4085	272	-1.95	0.0518	-0.7981	0.4085
Web only	-0.4711	0.1971	272	-2.39	0.0175	-0.4711	0.1971
DVD only	-0.3791	0.1304	272	-2.91	0.0040	-0.3791	0.1304
DVD + Web	-0.3471	0.2398	272	-1.45	0.1489	-0.3471	0.2398
All	-0.2819	0.1585	272	-1.78	0.0763	-0.2819	0.1585

Differences of Treatment Least Squares Means						
Treatment	_Treatment	Estimate	Standard Error	DF	t Value	Pr > |t|
Control	Web only	-0.3270	0.4422	272	-0.74	0.4603
Control	DVD only	-0.4190	0.4301	272	-0.97	0.3308
Control	DVD + Web	-0.4509	0.4931	272	-0.91	0.3613
Control	All	-0.5161	0.4244	272	-1.22	0.2250
Web only	DVD only	-0.09203	0.2376	272	-0.39	0.6988
Web only	DVD + Web	-0.1240	0.3264	272	-0.38	0.7043
Web only	All	-0.1892	0.2399	272	-0.79	0.4311
DVD only	DVD + Web	-0.03195	0.2710	272	-0.12	0.9062

Differences of Treatment Least Squares Means						
Treatment	**_Treatment**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
DVD only	All	-0.09713	0.2069	272	-0.47	0.6392
DVD + Web	All	-0.06518	0.3073	272	-0.21	0.8322

Confidence – *Cyberchase* math:

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
State	1	263	0.74	0.3904
Treatment	4	263	1.38	0.2408

Treatment Least Squares Means							
Treatment	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**	**Mean**	**Standard Error Mean**
Control	-0.3888	0.2518	263	-1.54	0.1239	-0.3888	0.2518
Web only	0.1311	0.1256	263	1.04	0.2973	0.1311	0.1256
DVD only	0.2025	0.08675	263	2.33	0.0203	0.2025	0.08675
DVD + Web	0.2689	0.1562	263	1.72	0.0863	0.2689	0.1562
All	0.1987	0.1033	263	1.92	0.0554	0.1987	0.1033

Differences of Treatment Least Squares Means						
Treatment	**_Treatment**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Control	Web only	-0.5199	0.2734	263	-1.90	0.0583
Control	DVD only	-0.5913	0.2675	263	-2.21	0.0279
Control	DVD + Web	-0.6576	0.3096	263	-2.12	0.0346
Control	All	-0.5875	0.2629	263	-2.23	0.0263
Web only	DVD only	-0.07140	0.1537	263	-0.46	0.6426
Web only	DVD + Web	-0.1377	0.2112	263	-0.65	0.5149
Web only	All	-0.06757	0.1540	263	-0.44	0.6611
DVD only	DVD + Web	-0.06633	0.1770	263	-0.37	0.7082
DVD only	All	0.003834	0.1362	263	0.03	0.9776
DVD + Web	All	0.07016	0.2000	263	0.35	0.7261

Confidence – non-*Cyberchase* math:

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
State	1	194	0.41	0.5229
Treatment	3	194	0.31	0.8197

Treatment Least Squares Means							
Treatment	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**	**Mean**	**Standard Error Mean**
Control	0.1048	0.3142	194	0.33	0.7391	0.1048	0.3142
Web only	0.2432	0.1745	194	1.39	0.1649	0.2432	0.1745
DVD only	0.2057	0.1399	194	1.47	0.1433	0.2057	0.1399
All	0.3391	0.1291	194	2.63	0.0093	0.3391	0.1291

Differences of Treatment Least Squares Means						
Treatment	**_Treatment**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Control	Web only	-0.1384	0.3388	194	-0.41	0.6833
Control	DVD only	-0.1009	0.3294	194	-0.31	0.7597
Control	All	-0.2343	0.3226	194	-0.73	0.4685
Web only	DVD only	0.03753	0.2062	194	0.18	0.8558
Web only	All	-0.09590	0.1960	194	-0.49	0.6252
DVD only	All	-0.1334	0.1741	194	-0.77	0.4445

Confidence – school math:

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
State	1	270	3.08	0.0804
Treatment	4	270	2.16	0.0737

Treatment Least Squares Means							
Treatment	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**	**Mean**	**Standard Error Mean**
Control	-0.4625	0.2932	270	-1.58	0.1159	-0.4625	0.2932
Web only	0.08190	0.1462	270	0.56	0.5758	0.08190	0.1462
DVD only	0.2794	0.09905	270	2.82	0.0051	0.2794	0.09905
DVD + Web	0.5458	0.1796	270	3.04	0.0026	0.5458	0.1796
All	0.1223	0.1199	270	1.02	0.3086	0.1223	0.1199

Differences of Treatment Least Squares Means						
Treatment	**_Treatment**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Control	Web only	-0.5444	0.3184	270	-1.71	0.0885
Control	DVD only	-0.7419	0.3111	270	-2.38	0.0178
Control	DVD + Web	-1.0083	0.3593	270	-2.81	0.0054
Control	All	-0.5848	0.3055	270	-1.91	0.0566
Web only	DVD only	-0.1975	0.1781	270	-1.11	0.2685
Web only	DVD + Web	-0.4639	0.2442	270	-1.90	0.0585
Web only	All	-0.04041	0.1785	270	-0.23	0.8211
DVD only	DVD + Web	-0.2664	0.2027	270	-1.31	0.1900
DVD only	All	0.1571	0.1576	270	1.00	0.3196
DVD + Web	All	0.4235	0.2317	270	1.83	0.0687

Confidence – non-math:

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
State	1	269	0.94	0.3332
Treatment	4	269	1.01	0.4019

Treatment Least Squares Means							
Treatment	Estimate	Standard Error	DF	t Value	Pr > |t|	Mean	Standard Error Mean
Control	-0.2589	0.3094	269	-0.84	0.4034	-0.2589	0.3094
Web only	-0.4179	0.1569	269	-2.66	0.0082	-0.4179	0.1569
DVD only	-0.2314	0.1034	269	-2.24	0.0260	-0.2314	0.1034
DVD + Web	-0.06051	0.1893	269	-0.32	0.7495	-0.06051	0.1893
All	-0.06778	0.1262	269	-0.54	0.5915	-0.06778	0.1262

Differences of Treatment Least Squares Means						
Treatment	_Treatment	Estimate	Standard Error	DF	t Value	Pr > |t|
Control	Web only	0.1589	0.3378	269	0.47	0.6383
Control	DVD only	-0.02753	0.3271	269	-0.08	0.9330
Control	DVD + Web	-0.1984	0.3786	269	-0.52	0.6007
Control	All	-0.1912	0.3230	269	-0.59	0.5545
Web only	DVD only	-0.1865	0.1887	269	-0.99	0.3240
Web only	DVD + Web	-0.3574	0.2583	269	-1.38	0.1676
Web only	All	-0.3501	0.1914	269	-1.83	0.0685
DVD only	DVD + Web	-0.1709	0.2144	269	-0.80	0.4260
DVD only	All	-0.1636	0.1642	269	-1.00	0.3199
DVD + Web	All	0.007263	0.2430	269	0.03	0.9762