MULTIMEDIA RESEARCH
33 BROWNS LANE • BELLPORT, NY 11713 • (631) 286-8925

Formative Evaluation of Stories for *CYBERCHASE*

Report for
WNET

by
Barbara N. Flagg
Multimedia Research

with assistance of

Harriet Davies
Pat Fisher
Ilona Holland

Research Report No. 99-003
February 18, 2000

TABLE OF CONTENTS

INTRODUCTION

With support from the National Science Foundation, WNET is developing an animated adventure series for 7-10 year olds. Entitled *Cyberchase,* the series involves a team of kids who use problem solving and logic to save the day on a dangerous and fun mission in cyberspace. This formative evaluation gathered feedback from third and fourth graders in response to two stories proposed for further development. The general goals for the formative evaluation were

- To assess reactions to the storylines, given two short audiotapes with color illustrations;
- To evaluate difficulty level and comprehension of the problem solving activities.

METHOD

Sample

In Milford, DE, two third grade and two fourth grade classes listened to two ten-minute audiotapes and viewed illustrative color drawings. Listeners included 39 third graders (61% female; 33% minority) and 41 fourth graders (56% female; 27% minority).

After listening, the whole class answered written appeal questions, while 12 students per class per show were interviewed individually. For "Lost My Marbles," 24 third graders were interviewed (58% female; 38% minority) and 24 fourth graders were interviewed (58% female; 42% minority). Because the classes did not contain at least 24 students, some students were interviewed again for the second show, "Snow Day." For "Snow Day," the 24 interviewed third graders included 54% female and 42% minority and the 24 fourth graders included 54% female and 21% minority.

Procedure

Four researchers were present in the classrooms to show the pictures and interview students. The head researcher introduced the series concept and procedure. To familiarize respondents with the characters and animation style, students viewed the opening two minutes of the pilot show of *Cyberchase*. The whole class then listened to the audiotape and viewed pictures for "Lost My Marbles" and responded to written appeal questions. Each of the four researchers then interviewed three students individually for seven minutes about content issues. The same procedure was repeated for "Snow Day."

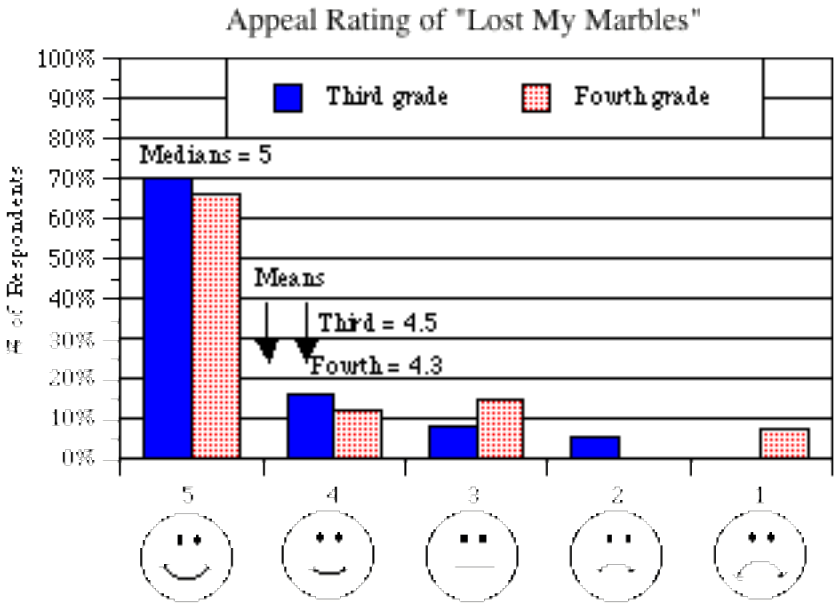
All data were examined with respect to gender, grade and majority-minority group differences. Differences beyond a chance result ($p < .05$) are reported as "significant." Percentages in tables are rounded off to the nearest whole number.

LOST MY MARBLES: APPEAL

Lost My Marbles appealed to 87% of third graders and 78% of fourth graders; 68% of the whole sample liked the show very much. The third graders liked the show significantly more than the fourth graders. Gender did not influence appeal whereas ethnicity did only within third grade - white students rated the show higher in appeal than did minority students.
About three-quarters of both grades felt that the producers should go ahead to make the show.

When asked what they liked about the show, one-fifth of students focused on the video piece. In response to the audiotape, the audience liked the broccoli pit reference (9%), the successful solution to the story (9%) and that the kids used math or used their brains (9%). Fewer viewers could identify what they did not like about the show, with 39% writing that they liked everything. Small groups of children focused on not liking the bad guys (11%); the broccoli pit (5%) or not liking anything (5%).

Appeal of the Show



After listening to the show, 87% of third graders and 78% of fourth graders were positive in their response about how much they liked the show, circling a "very happy" face or a "sort of" happy face (see chart). Translating the face scale to numbers (5 to 1) yields an overall mean appeal rating of 4.5 out of 5 for third graders and 4.3 for fourth graders.[1]

The third grade students rated the show higher in appeal than fourth grade students ($p \leq .05$; Mann-Whitney U test of medians). Within third grade, there were no differences for gender but significant differences for ethnicity; third grade white students liked the show more than minorities ($p \leq .05$; Mann-Whitney U). Within fourth grade, there were no differences for gender or ethnicity.

[1] In response to the pilot video program, a 3rd/4th grade sample gave a mean rating of 4.8 out of 5, with no grade differences.

Students' Recommendation to Producers

Students were asked if the producers should go ahead to make a television show out of the story. About three-quarters of both age groups answered "yes." (see table below):

	Yes	**Maybe**	**No**
% **3rd** (N=39)	74%	23%	3%
% **4th** (N=41)	78%	15%	7%

These results were not influenced by grade, gender or ethnicity. The results were related to appeal ratings. Those who said "yes, make the show" produced a mean appeal rating of 4.8 (n = 61), significantly higher than those who said "maybe" (mean = 3.4; n = 15) and those who said "no" (mean = 2.5; n = 4).

What Students Liked and Did Not Like about the Show

The students were asked to write down what they liked and did not like about the show and why. In defining what they liked best, 19% responded about the video rather than the audio show (e.g., "liked when Jackie was sucked into the computer"). In defining what they did not like, 39% responded that they liked everything. The top categories of remaining responses appear below:

Top categories of "liked best" responses:

- 9% like broccoli pit because funny
- 9% liked that the kids found Marbles and could fix mother board
- 9% liked that they used math; made numbers/lines on map; used brains; were smart
- 6% liked how characters looked - drawn; animated; weird-liked it
- 6% liked the action, adventures, excitement

Top categories of "did not like" responses:

- 11% didn't like the 'bad guys' because they were bad guys: "stupid," "weak," "trying to stop the kids," "mean," "evil", "let virus go," "put smoke out"
- 5% did not like the broccoli pit because "not like broccoli" "smells nasty" "yucky"
- 5% did not like anything, boring, not cool

Comments for Scripting

Overall, this show appears to be competitive with the pilot show in terms of general appeal. Students seemed to like the basic concept of an adventurous team using their brains against the 'bad guys.'

Examination of the "like/dislike" answers of the third grade minorities who rated the show lower than the white students did not reveal any explanations. However, the interviews about problem solving behaviors may shed light on the appeal ratings (see pgs. 6-12). The third grade minorities had larger representations in the category of "lacking knowledge" for all three story problems, as indicated in the following table:

Percent of Students Lacking Knowledge in the Three Story Problems

Story Problem	Third grade minority	Third grade white	Fourth grade minority	Fourth grade white
Map/Landmark	**56%**	**27%**	20%	21%
Scale/Distance	**44%**	**13%**	30%	7%
Map/Grid	**78%**	**60%**	30%	21%

In addition, 46% of minority third grade students rated the problems as "hard" compared with 19% of white third grade students. Note that these differences between majority and minority third graders, although seemingly large, are not statistically significant, probably because of our small samples. The differences in problem understanding and difficulty ratings simply indicate one possible explanation for appeal differences between these groups.

LOST MY MARBLES: AGE APPROPRIATENESS

In choosing who would watch the series, at least three-quarters of each grade felt that the show was appropriate for kids their own age; slightly over half of each grade felt it was also for younger kids but only 21% of third graders and 7% of fourth graders thought that older kids would watch the show. Responses about which age group would watch appeared to be related to appeal of the show and perceived difficulty of the mapping problems.

About three-quarters of the sample felt the mapping problems were "just right" in difficulty and one-fifth felt they were "easy." Boys felt that the problems were easier than girls.

Respondents were asked who will watch this television show on TV. Students could circle any or all of three responses (see table below). There were no gender, grade, or ethnicity differences in these results. Responses related to students' rating of show appeal and problem difficulty.

Who will watch this television show on TV?

For comparison, %s in parentheses are responses from pilot video evaluation

	N =	Kids younger than me	Kids my age	Kids older than me
Third grade	39	56% (63%)	82% (85%)	20% (23%)
Fourth grade	41	56% (74%)	76% (72%)	7% (35%)
All	80	56% (68%)	79% (79%)	14% (29%)

More than three-quarters of students in both the third and fourth grades felt the show was for kids their **own age**; this is a response percentage equivalent to that found in the evaluation of the pilot video (*Poddleville*). Those who felt the show was appropriate for them rated the show higher in appeal (mean = 4.7) than those who did not choose this response (mean = 3.4). Of those who felt the show was for kids their own age, 78% rated the mapping problems as "just right" in difficulty. In contrast, of those who did not agree that the show was for their age, 47% rated the problems as "just right."

Slightly over half of each grade felt that the show was appropriate for **younger** kids. These students rated the show lower in appeal (mean = 4.1) than those who did not choose this response (mean = 4.7). Of those who felt the show was okay for younger kids, 29% thought the mapping problems were "easy," whereas of those who did not agree that the series was for younger kids, 11% thought the problems were "easy."

Finally, 21% of third graders and 7% of fourth graders thought that **older** kids would watch the show. These students rated the show higher in appeal (mean = 5.0) than those who did not choose this response (mean = 4.3). Problem difficulty did not relate significantly to the "older" choice.

As indicated above, students were asked to rate the difficulty of the mapping problems (see table below). About three-quarters of the sample felt the problems were "just right"

and one-fifth felt they were "easy."

Do you think that the mapping problems were hard, easy or just right for kids your age?

	N =	Hard	Just right	Easy
Third grade	39	8%	74%	18%
Fourth grade	41	7%	68%	24%
All	80	8%	71%	21%

These results were not influenced by grade or ethnicity but gender showed an effect overall. Boys felt that the problems were easier than girls (Chi-square significant at $p = .004$)

	N =	Hard	Just right	Easy
Girls	47	9%	83%	9%
Boys	33	6%	55%	39%

Comments for Scripting

This show appears to play better with third graders. Third graders rated the show appeal significantly higher than fourth graders (p. 2), and more third graders (not significant) rated the show as appropriate for their age and older (p. 4).

PROBLEM COMPREHENSION: LOST MY MARBLES

Map/Landmark Problem

With visual and verbal prompting, 71% of students could explain the map/landmark relationship for Skull Cliff. Without prompting, 46% of students demonstrated substantial knowledge. There were no significant grade, gender or ethnicity relationships in these interviews.

The 24 interviewed students were asked the following question without reference to an illustration:

When the kids landed on the island, they had a map but they didn't know where they were. How did they figure out where they were on the map? Anything else?

The expected answer would indicate that the kids went to the ocean or edge of the map and/or found a landmark or skull cliff. Students were then asked:

After the kids reached the ocean and the beach, how did they figure out where they were on the map?

Respondents were expected to say they saw a landmark/skull cliff that was marked on map. Students were then shown an illustration of Jackie holding the map and seeing Skull Cliff in the distance, and they were asked:

What kind of landmark did the kids see?

How did seeing Skull Cliff help them figure out where they were on the map?

Without verbal or visual prompting, almost half (46%) of the students demonstrated substantial knowledge of how the kids located themselves on the map, finding a landmark that matched up with the map. About half of these students also recalled that the kids first went to the ocean or beach. With the verbal prompt question, another 10% of the sample, all fourth graders, described the relationship between the cliff and the map landmark. With the addition of the illustration and two prompt questions, 15% more of the sample managed to make the connection between the skull on the land and that on the map.

Thus, 63% of the third graders and 79% of the fourth graders could explain the map/landmark relationship for Skull Cliff, although some students needed both visual and verbal prompting. There were no significant grade, gender or ethnicity relationships in these interviews.

The table on the following page presents more detail of the categories used to code the interview responses. "Substantial knowledge" means that the students explained that the kids located themselves on the map by seeing the cliff in front of them and locating the cliff on the map.

Response Type	Example Answer	3rd	4th	All
Substantial knowledge without prompt	They walked along the beach until they came to a landmark and they matched it up with a picture on the map to see if they were the same. They found a landmark, a big rock like a skull, on the land and on the map.	50%	42%	46%
Substantial knowledge with question prompt	(without) They put twigs where x was, how far to go to the guy they were looking for. (w/prompt) They pulled out the map and saw the same rock on the island.	0%	21%	10%
Substantial knowledge with visual and question prompts	(without) They looked on the map. (w/question prompt) I don't know. (w/visual and questions) They saw a skull, and they looked on the map and seen the skull too.	13%	17%	15%
Substantial knowledge		**63%**	**79%**	**71%**
Lacking knowledge without & with visual and question prompts	(without) They did the ABC and the 123. (w/question prompt) I don't think they did. They just guessed it [where they were]. (w/visual and questions) The rock there with the skull figure. . . So they could see how many paces to get to the mark.	38%	21%	29%
Lacking knowledge		**38%**	**21%**	**29%**

Comments for Scripting

The piece of information missing in most responses was the initial step that the kids went to the ocean/edge of the map and why. From our data, it is impossible to determine whether this is an artifact of the interview procedure or a real scripting problem.

Scale/Distance Problem

With visual and verbal prompting, 34% of the sample could recall three or four of the four steps Jackie took to solve the problem of how far they had to go to reach Marbles, and 46% of the sample could recall one or two of the four steps. The remaining 23% of students were not able to describe coherently any of the steps. There were no significant grade, gender or ethnicity relationships in these interviews.

Most students had difficulty recalling that Jackie counted paces from the beach to the cliff; only 29% mentioned this step. 67% understood that the twig stood for a distance; 58% mentioned counting twigs; and 47% recalled some multiplication to calculate distance.

The 24 interviewed students were asked the following question without reference to an illustration:

Jackie drew a line on the map from Skull Cliff to the x where Marbles was. How did the kids figure out how far they had to walk to find Marbles? ... Anything else?

The expected answer would indicate that Jackie counted paces from the beach to the cliff, used a twig to stand for the number of paces, counted the number of twigs that fit on the map between the beach and the x and finally multiplied the number of twigs by paces from the beach to the cliff. Students were then shown illustrations of Jackie walking to the skull landmark and Jackie holding the twig on the map, and they were asked the following:

(1) Why did Jackie count her steps from the edge of the beach to Skull Cliff?

(2) What did the twig stand for or represent on the map?

(3) How did Jackie use the twig to measure the distance to where Marbles was?

The expected answers were (1) to find out the distance or how big distances were on the map; (2) the twig meant 200 steps or a certain distance; (3) counted eight twig lengths from the skull to the x, then multiplied 8 x 200 paces equaling 1600 paces total distance.

The table below presents the distribution of responses where each student response was coded for one of five possibilities:

"Substantial knowledge" is defined for this set of questions as mentioning three or all four steps that Jackie took, although the numerical detail might be incorrect:

(1) Establish distance scale: Jackie counted paces from beach to cliff.

(2) Twig as scale symbol: She used a twig to stand for number of paces.

(3) Scale lengths: She counted how many twigs fit between the beach and the x.

(4) Calculate distance: Multiplied number of twigs by paces from beach to cliff.

"Partial knowledge" for this set of questions means that the student mentioned only one or two of the steps, without or with the prompt illustrations and questions.

Response Type	Example Answer	3rd	4th	All
Substantial knowledge without prompt	Jackie paced to Skull Cliff-200 paces and she said one twig is 200 paces. Then she got more twigs and counted how many she needed to get to the x. . . then she said, 8 times 200 equals 1600.	8%	25%	17%
Substantial knowledge with visuals & verbal prompts	(w/out) Jackie figured out that a little twig is about 200 centimeters. Then she realize it was about 200' on the map. [count steps?] Because she was counting how many steps to get there. [twig stand for?] 200'. [how use twig?] by counting 200 times 7 twigs. She took a bunch of twigs and put them on the map and then she was able to see how many it would take. 200 x 7.	13%	21%	17%
Substantial knowledge		**21%**	**46%**	**34%**
Partial knowledge without prompts (prompts elicited no new info)	(w/out) They used a twig. They found out how much one twig was and then they marked it all the way to the x.	8%	21%	15%
Partial knowledge with visual/verbal prompts	(w/out) They counted. They did 8 x 200. [count steps?] So she's know it'd be a straight line. She'd know steps to there. [twig stand for?] 200 paces. [how use twig?] She put it on the map.	46%	17%	31%
Partial knowledge		**54%**	**38%**	**46%**
Lacking knowledge with-out & with prompts	(w/out) They did ABC and 123 and went to D4. [count steps?] To get the woman. [twig stand for?] To make the numbers and letters. They moved the twig to the x. [how use twig?] I don't know.	29%	17%	23%
Lacking knowledge		**29%**	**17%**	**23%**

To summarize the table results, few third graders (8%) but one-quarter (25%) of the fourth graders were able to recall without prompting three or four of the steps that Jackie took. With prompting, 21% of third graders and 46% of fourth graders showed "substantial knowledge" of this problem and solution. So, about one-third of the whole sample demonstrated "substantial knowledge."

"Partial knowledge," that is, recall of one or two of the steps without or with prompts, was demonstrated by half (54%) of the third graders and one-third (38%) of the fourth graders.

Those unable to explain coherently any of the four steps taken in the problem solution included 29% third graders and 17% fourth graders.

In the table below, we examine the responses by looking at how many students mentioned each of the four steps in their whole interview. More than half of the students could recall that the twig stood for a certain distance and that Jackie counted twig lengths along the line. Almost half recalled the calculating of the distance, although some were incorrect in their figures. Only 29% recalled the beginning effort to establish the distance scale by pacing off from beach to cliff. There were no grade or gender differences in these data; however, for the fourth grade, minority students mentioned a mean of 1.3 steps, significantly less than white students at a mean of 2.8 steps out of the 4 possible. The minority students had more difficulty recalling the counting of paces and the calculating of distance.

Solution Steps	3rd	4th	All
Establish distance scale: Jackie counted paces from beach to cliff	25%	33%	29%
Twig as scale marker: Jackie used twig to stand for number of paces or certain distance	42%	67%	67%
Scale lengths: Jackie counted how many twigs fit between the beach and the x	58%	58%	58%
Calculate distance: Jackie multiplied number of twigs by paces from beach to cliff	35%	58%	47%

Comments for Scripting

The fact that Jackie counted off the paces to establish a distance scale was not clear to most students. A good portion of students understood that the twig was to represent a certain distance but many did not connect it to Jackie's earlier "pacing off the distance." Perhaps flipping the order would help: Jackie chooses a twig and breaks it off so it matches the space on the map between the beach and cliff. Jackie then notes that she has to count off the paces between the beach and the cliff to know how many steps the twig represents - and so forth. This order appears to be more aligned with what the students say naturally, "She had a twig. She counted 200 paces;" "They used a twig and marked how many paces they would have to walk. She took the paces and multiplied it by the number of twigs." "They used a leaf. They counted how many steps from the stone;" "They used a twig. They found out how much one twig was ...." In the open-ended answers, the concept of the scale marker frequently came out first followed by the distance represented by that marker.

Some students also were unfamiliar with the measurement system of "paces," suggesting instead traditional measurement by "centimeters," "feet," "meters," "miles," "yards," and "inches."

Some imagined that Jackie used eight different twigs rather than moving the same twig eight times: "she got more twigs;" "used twigs and set them on the line;" "she put one twig down and marked the spot where it ended and then put another twig down." The animation can clarify how the same twig is used repeatedly or the kids can make eight twigs of the same length and lay them out on the line.

Map Grid Problem

With the visual prompt of the map illustration, 100% of the fourth graders and 92% of the third graders could use their fingers to locate the bridge on the map, and 92% of both the third and fourth graders could locate and identify the pond at B6. However, when queried without an illustration about how the kids met up in the fog, many students had difficulty. A complete explanation about the map grid problem was provided by no third graders and 29% of fourth graders. About 40% of the sample gave a partial answer focusing on either the grid, the letters and numbers, or the intersection. Those unable to explain coherently any of the steps included 67% of third graders and 25% of fourth graders. Third grade females showed the most difficulty with the uncued interview question.

The 24 interviewed students were asked the following question without reference to an illustration:

The kids got separated in the fog. How did they use the map to find out where they were and get together again?

The expected answer would mention three steps: Label one side of grid with letters; Label other side of grid with numbers; Agree to meet at an intersection of a letter and number. Students were then shown an illustration of the map with grids, letters and numbers and the x and asked the following:

Show me how the kids used the letters and numbers to find the bridge on the map? Can you tell me what is at B6?

With the visual prompt of the map illustration, 100% of the fourth graders and 92% of the third graders could use their fingers correctly to locate the bridge on the map at D4 or locate an intersecting letter and number, if they could not recall the D4 specification. In response to the same illustration, 92% of both the third and fourth graders could locate and identify the pond at B6. The two third graders had no idea how to use the grid, and one fourth grader named all landmarks on the 6 grid line while the other fourth grader went up the B grid line to 6 but then counted to the right another 6 grid blocks, ending up at G6. Thus, students were very capable of using the letter/number grid system to locate positions; however, they were less able to explain how the kids located themselves in the fog by generating the grid system, as explained below.

The table below presents the distribution of responses where each student response was coded for one of five possibilities:

"Substantial knowledge without prompt" is defined as mentioning the labeling by letters and numbers and meeting at an intersection of a letter and number.

"Partial knowledge without prompt" included (a) those who labeled with letters and numbers but did not explain choosing an intersection; (b) those who described making a grid and choosing a meeting place but did not mention letters or numbers; and (c) those who described meeting at a letter/number intersection but not the process of marking the map.

"Lacking knowledge" means that students simply said the kids agreed on a place to meet or gave an entirely irrelevant answer.

Response Type	Example Answer	3rd	4th	All
Substantial knowledge without prompt	They used cyberpads and copied the map. They put letters and numbers on the map. They met at D4 at the bridge.	0%	29%	15%
Substantial knowledge		**0%**	**29%**	**15%**
Partial knowledge without prompts	(a) They put letters and number on the map.	17%	13%	15%
	(b) They turned the map into a grid. Jackie sent a copy. Then they decided to meet on the bridge.	8%	21%	15%
	(c) They picked a certain place, D4, and they went to the same place because they knew where they were going.	8%	13%	10%
Partial knowledge		**33%**	**46%**	**40%**
Lacking knowledge with-out prompts	They came up with a place they should meet and they used the map.	67%	25%	46%
Lacking knowledge		**67%**	**25%**	**46%**

To summarize the table results, no third graders but one-quarter (29%) of the fourth graders were able to provide a complete explanation, including labeling by letter and number and meeting at an intersection. Some fourth graders gave the correct D4 meeting place; some mentioned another intersection from the show like A3; and a few also explained the sharing of the map.

About 40% of the sample provided a partial answer focusing on either the grid, the letters and numbers or the intersection.

Those unable to explain coherently any of the steps taken in the problem solution included 67% of the third graders and 25% of the fourth graders. In this category were 50% of the third grade females, a significant portion.

Three fourth graders (6% of sample) demonstrated "substantial knowledge" for all three of the show problems (map/landmark; scale/distance; grid). Four third graders (8% of sample) "lacked knowledge" for all three problems.

Comments for scripting

Students were able to use the map grid, letters and numbers, to locate a known and unknown landmark, which indicates familiarity with a longitude/latitude system; some students even used these terms. However, 67% of third graders found it difficult to explain the steps that the kids used to get together, even with minimal prompts. It's possible that with more individualized probing, we would have improved the success rate; on the other hand, by this point in the show, the listeners had been exposed to a high number of problem-solving strategies and the younger ones may simply have lost the story thread in the multi-step complexity.

Note also that girls felt the mapping problems were harder than boys did (p. 4) and that third grade girls had more difficulyy than their male cohorts explaining the grid problem

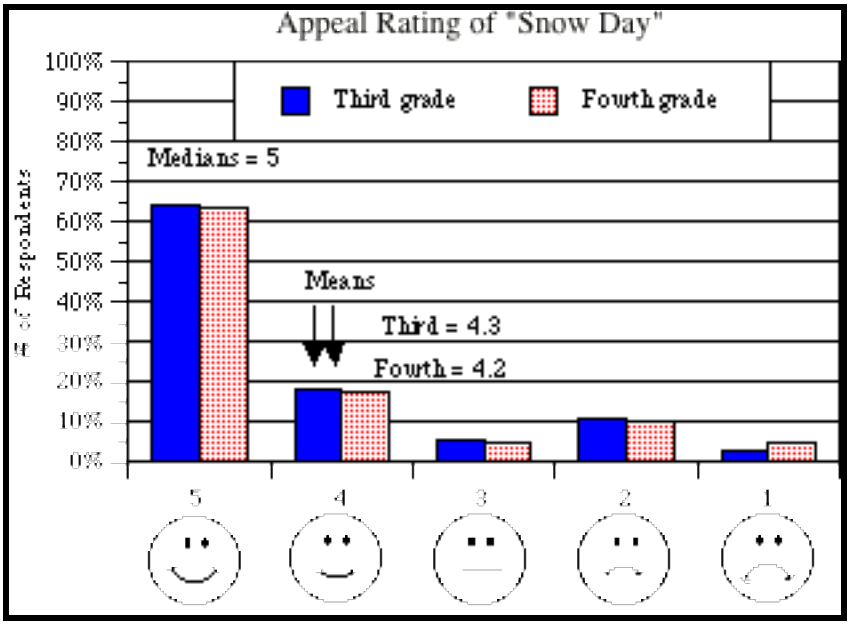
(p. 12). Spatial-visualization is one of the few math skills in which gender differences continue to hold up in the research literature.

SNOW DAY: APPEAL

Snow Day appealed to 82% of third graders and 81% of fourth graders; 64% of the whole sample liked the show very much. Those who felt that the producers should go ahead to make the show comprised 56% of third graders and 71% of fourth graders.

When asked what they liked about the show, one-fifth of respondents focused on the seal story; 15% on the estimation problem solving strategy; 13% on the barrels story; 10% on Inez shooting the sunisphere like a basketball; and 8% on the happy ending. Fewer viewers could identify what they did not like about the show, with 40% writing that they liked everything. About 14% reported not liking the bad guys.

Appeal of the Show



After listening to the show, 82% of third graders and 81% of fourth graders were positive in their response about how much they liked the show, circling a "very happy" face or a "sort of" happy face (see chart). Translating the face scale to numbers (5 to 1) yields an overall mean appeal rating of 4.3 out of 5 for third graders and 4.2 for fourth graders.[2] There were no differences between grades in appeal ratings. Within grades, there were no differences for gender or ethnicity.

[2] In response to the pilot video program, a 3rd/4th grade sample gave a mean rating of 4.8 out of 5, with no grade differences.

Students' Recommendation to Producers

Students were asked if the producers should go ahead to make a television show out of the story. More fourth graders said "yes" than third graders, but the distribution of answers was not influenced significantly by grade, gender or ethnicity (see table below):

	Yes	**Maybe**	**No**
% **3rd** (N=39)	56%	33%	10%
% **4th** (N=41)	71%	22%	7%

The results were related to appeal ratings. Those who said "yes, make the show" produced a mean appeal rating of 4.9 (n = 51), significantly higher than those who said "maybe" (mean = 3.6; n = 22) and those who said "no" (mean = 1.7; n = 7).

What Students Liked and Did Not Like about the Show

The students were asked to write down what they liked and did not like about the show and why. In defining what they did not like, 40% responded that they liked everything. The top categories of responses appear below:

Top categories of "liked best" responses:

- 20% liked the story with the seals because they "like seals;" it was a "fun math problem;" or just "funny."
- 15% liked the fact that the kids estimated because it's a "good fast strategy" and "it's "quicker."
- 13% liked the barrels story because it was a "good idea;" they "estimated;" "they were going fast;" "they bumped."
- 10% liked that Inez shot the sunisphere like a basketball because "she believed in herself;" "like basketball/sports."
- 8% liked the happy ending that the kids saved the town.

Top categories of "did not like" responses:

- 14% didn't like the 'bad guys' because they were bad guys: "stealing is wrong;" "greedy;" "town could have frozen."

Comments for Scripting

Overall, this show appears to be competitive with the pilot show in terms of general appeal level. Students liked the contrast of estimating and measuring, and they enjoyed the fun of the seals playing keepaway, the bumpy ride on the barrels and Inez shooting the sunisphere as a basketball. These aspects of fun kept their interest and lightened the learning without distracting them from the teaching moments. Note that not one student said they didn't like the show because of the math content; but this was a school setting and videos obtain a halo effect.

SNOW DAY: AGE APPROPRIATENESS

In choosing who would watch the series, about three-quarters of the sample felt that the show was appropriate for kids their own age, about half felt it was also for younger kids but only 19% thought that older kids would watch the show. Responses about which age group would watch appeared to be somewhat related to appeal of the show and perceived difficulty of the estimation problems.

Almost two-thirds of the sample felt the estimation problems were "just right" in difficulty and almost one-third felt they were "easy." Boys felt that the problems were easier than girls.

Respondents were asked who will watch this television show on TV. Students could circle any or all of three responses (see table below). There were no gender, grade, or ethnicity differences in these results.

Who will watch this television show on TV?

For comparison, %s in parentheses are responses from pilot video evaluation

	N =	Kids younger than me	Kids my age	Kids older than me
Third grade	39	44% (63%)	79% (85%)	23% (23%)
Fourth grade	41	51% (74%)	71% (72%)	15% (35%)
All	80	48% (68%)	75% (79%)	19% (29%)

About three-quarters of students felt the show was for kids their **own age**; this is a response percentage equivalent to that found in the evaluation of the pilot video (*Poddleville*). Those who felt the show was appropriate for them rated the show higher in appeal (mean = 4.7) than those who did not choose this response (mean = 3.2). Of those who felt the show was for kids their own age, 72% rated the estimation problems as "just right" in difficulty. In contrast, of those who did not agree that the show was for their age, 35% rated the problems as "just right."

About half of the respondents felt that the show was appropriate for **younger** kids. These students rated the show lower in appeal (mean = 3.8) than those who did not choose this response (mean = 4.7). Problem difficulty did not relate significantly to "younger" choice.

Finally, about one-fifth of the students thought that **older** kids would watch the show. Appeal and problem difficulty did not relate significantly to the "older" choice.

Students were asked to rate the difficulty of the estimation problems (see table below). Almost two-thirds of the sample felt the problems were "just right" and almost one-third felt they were "easy."

Do you think that the estimation problems were hard, easy or just right for kids your age?

	N =	Hard	Just right	Easy
Third grade	39	10%	62%	28%
Fourth grade	41	5%	63%	32%
All	80	8%	63%	30%

These results were not influenced by grade or ethnicity but gender showed an effect overall. Boys felt that the problems were easier than girls (Chi-square; $p = .04$)

	N =	Hard	Just right	Easy
Girls	47	9%	72%	19%
Boys	33	6%	49%	46%

Comments for Scripting

Most students felt that this show was meant for them in terms of appeal and problem difficulty. However, one might generalize from these data to conclude that fifth graders, particularly boys, will find the fully animated show on the young side and the content problems on the easy side. Including action and comic action by Matt and Hacker (such as the snowboarding and throwing of snowballs) will maintain interest of the older boys.

PROBLEM COMPREHENSION: SNOW DAY

Differences in Problem Solving Strategy

Two-thirds of the fourth graders and three-quarters of the third graders could explain satisfactorily the differences in problem solving strategies among the characters; that is, that Matt and Jackie estimated and that Inez measured exactly. Minority children had somewhat more difficulty explaining the difference in approach, although the data are not clear as to why that was the case.

Students were showed an illustration of the three characters with their names and asked:

The kids tried to solve problems differently in the beginning of the story. What is the difference between the way Inez tried to solve problems and the way Matt and Jackie tried to solve problems?

Two-thirds (67%) of the fourth graders and three-quarters (79%) of the third graders recognized that Matt and Jackie estimated or guessed and that Inez measured or counted to be exact. Of those not able to explain the differences, 62% were minority children. Correct answers included, for example:

Inez didn't like to estimate, she liked to get the exact answer. Matt and Jackie liked to estimate. (WF4)

Matt and Jackie estimated. Inez didn't want to estimate. She wanted to do it right. (WM4)

Matt and Jackie liked to estimate. Got there a lot quicker. Inez liked to make right measurements. (WF3)

Matt tried to estimate. Estimation worked. Inez tried to measure. Time could have been out if you measured exactly. (WM3)

Comments for Scripting

Note that a number of students could use the term "estimate" but their later explanations of how specific problems were solved indicated that they parroted back the vocabulary but did not truly understand what "estimation" means. A few students, like some of those quoted above, noted that estimation made measurement quicker, but others, also quoted above, implied that estimation was a 'not right' measurement compared with Inez's "right measurements."

Bridge Problem

About one-quarter of the students demonstrated substantial knowledge of the bridge problem and its solution, recalling both the information about the bottom layer and the attempt to estimate with Digit's wing. Another 46% of the sample demonstrated partial knowledge, describing either the bottom layer information or the estimation, but typically the latter piece of the solution. One-third of the sample were not able to recount the bridge problem/ solution with any coherence. Gender and grade did not play a significant role in the interview responses.

The 24 interviewed students were asked the following question without reference to an illustration:

> **In pursuit of the golden ball, the kids could not cross the ice bridge because there was a huge hole in the middle. They decided to use barrels to fill the hole. How did Matt and Jackie figure out how many barrels they needed to fill the hole?**

The expected answer would indicate that the kids put in four barrels to make a bottom layer, then used Digit's wings to estimate how many more barrels they needed. Students were then shown an illustration of the ice bridge with Digit next to the single layer of barrels. Students were asked the following:

> **How did Digit help the kids figure out how many more layers of barrels they needed to fill the hole?**

The expected answer would suggest that Digit measured the size of the hole with his wings like a ruler. One wingspan is the width of barrel.

About one-quarter (23%) of the students demonstrated substantial knowledge of the bridge problem, recalling both the information about the bottom layer and the attempt to estimate with Digit's wing. Another 46% demonstrated partial knowledge, describing either the bottom layer information or the estimation, but typically the latter piece of the solution. One-third of the sample were not able to recount the bridge problem/solution with any coherence. Gender and grade did not play a significant role in the interview responses. More detail is provided in the following table.

In the following table, student responses to the above question set were coded for one of five response types:

"Substantial knowledge" included both the bottom layer information and estimation with Digit's wing.

"Substantial knowledge with prompt" where prompt refers to the illustration and question.

"Partial knowledge" included either bottom layer information or estimation, but not both.

"Partial knowledge with prompt"

"Lacking knowledge" means no coherent answer was given including the above information.

Response Type	Example Answer	3rd	4th	All
Substantial knowledge without prompt	They estimated by taking four barrels and putting them on the bottom. Then they used the bird. His wing was the width of the barrels. They counted his wings to the top of the hole to find how many more layers of barrels they needed.	8%	13%	10%
Substantial knowledge with visual & question prompt	(w/out) First they put 2 and 2 in the hole and then they estimated and predicted how many they would need to get up. The answer was 16. (with) He took his wing and measured how long it would be.	8%	17%	13%
Substantial knowledge		**17%**	**29%**	**23%**
Partial knowledge without prompt (prompt elicited no new info)	The bird put out his wings and measured like 4 per layer. They used the bug's wings to measure how many more layers they needed.	17%	29%	23%
Partial knowledge with prompt	(w/out) They estimated and if they needed more they could get more. (with) Digit used his wing and the barrel was about Digit's wing so they did it by estimation with Digit's wing.	33%	13%	23%
Partial knowledge		**50%**	**42%**	**46%**
Lacking knowledge without & with prompt	(w/out) They estimated and used times tables (with) I don't know. (w/out) By division. (with) By his wings. (w/out) By counting up. (with) She counted up by the wall, counted the barrels.	33%	29%	31%
Lacking knowledge		**33%**	**29%**	**31%**

Comments for Scripting

The main portion of the solution that was missed in students' recounting was the reasoning behind the initial four barrel layer. One-third of the students mentioned putting down the first layer, but only one student appeared to understand that a base layer was needed to estimate the size of the hole before estimating number of layers.

About 70% of the students understood that Digit used his wings to "measure" or "estimate" how many extra layers of barrels were needed. However, many students used the words "measure" and "estimate" here interchangeably, as did the script; this lack of distinction may add to confusion about what "estimation" is compared to exact "measurement." Also, students were not clear on whether Digit was measuring the width or height of one barrel or a whole layer; respondents were clear on the general idea of using Digit's wing but not on the specific use. Since width of a round item is difficult to measure/estimate with a wing or otherwise and since the barrels would actually roll out of the hole as drawn, consider repositioning the barrels on their bottoms, as upright cylinders, then Digit's wing can match one barrel's height.

Seal and Fish Problem

About one-third of the students demonstrated substantial knowledge of the seal and fish problem and its solution, recalling information about counting seals on one ice piece, adding another to the count and multiplying by the number of ice pieces. Another 56% of the sample demonstrated partial knowledge, describing one or more of the three steps; typically, students left out the addition of another seal/fish for a safety margin. Only 8% of the sample were unable to explain clearly any part of the problem and solution. Gender and grade did not play a significant role in the interview responses.

The 24 interviewed students were asked the following question without reference to an illustration:

Inez decided to get the golden ball from the seals by feeding them fish. The kids thought that if they could feed all the seals at once, they could get the ball. How did they figure out how many seals there were?

The expected answer would indicate that the kids estimated the number of seals by counting how many seals were on one ice piece and multiplying by 6 pieces. Students were then asked:

How did they make sure they had enough fish?

The team added one more seal per piece of ice to make an estimate of 9 seals per piece. Students were then show an illustration of the six ice pieces with seals on each. Students were asked the following:

Use this picture to show me how Matt and Jackie figured out how many seals there were.

About one-third (35%) of the students demonstrated substantial knowledge of the seal problem, recalling the three steps of counting seals on one piece, adding one more seal and multiplying by number of ice pieces. Another 56% demonstrated partial knowledge that included one or two of the three steps. Typically, the missing step was an understanding that an additional seal was included in the estimate per ice piece. Only 8% of the sample was not able to recount the seal problem and solution with any coherence. Gender and grade did not play a significant role in the interview responses. The following table provides a more detailed analysis.

In the following table, student responses to the above question set were coded for one of five response types:

"Substantial knowledge" included the counting of seals on one piece, the addition of one more seal and the multiplication by number of ice pieces.

"Partial knowledge" included one or two of the three steps in the "substantial" definition.

"Lacking knowledge" means no coherent answer was given.

Response Type	Example Answer	3rd	4th	All
Substantial knowledge without prompts	By counting the seals on one block and adding one. In their heads they did the math. They said 9 times 6 to find the answer. There were 6 pieces of ice and Matt said 'let's estimate.' There's 8 on each piece. Inez said let's say 9 to be safe. They timesed 6 times 9 equals 54.	25%	38%	31%
Substantial knowledge with question prompt	They multiplied 6 times 9. There were 6 ice sheets and 9 seals on the ice. (w/question) They estimated then added one more to every sheet.	-	8%	4%
Substantial knowledge with visual prompt		-	-	-
Substantial knowledge		**25%**	**46%**	**35%**
Partial knowledge without prompts (prompts elicited no new info)	They multiplied the number of ices by 8. [how get 8?] They counted one block of seals. It equaled 8 and then multiplied by the number of ices.	29%	25%	27%
Partial knowledge with prompts	(w/out) They added one more to each one and there was 9 and there was 56. (w/question) They added one more. (w/visual) They imagined there was one more on each ice and times the pieces of ice they had. (w/out) They estimated. (w/question) I don't know. (w/visual) They figured about the same number of seals on each piece of ice and they added it up. (w/out) I forgot that part. (w/question) By going fishing and getting 54 fish to feed each seal. (w/visual) They added one more seal to each. There was 8 [points to ice floe] and they put one more on each and estimated how many they'd need. [then?] I can't remember.	33%	25%	29%
Partial knowledge		**63%**	**50%**	**56%**
Lacking knowledge without & with prompts	(w/out) By division. They gave every seal one fish. (w/question) They got one for every seal that was on the ice. (w/visual) They used 9 on every one. So they did 9 divided by 6. They made sure they got the exact number so they could feed every seal.	12%	4%	8%
Lacking knowledge		**12%**	**4%**	**8%**

Comments for Scripting

Students knew that the kids had counted eight seals on a piece of ice and many used the word "estimate," but many students' responses implied that they thought every piece of ice had eight seals (e.g., "there's eight on each piece"). Only one student's response showed an understanding that some pieces would have more seals and some pieces would have less and that eight was to stand for an average number or a guess at what might be on all the pieces. Clarify the initial process of estimating.

Many students did not recall the addition of a seal/fish to the eight seals counted, although those who did understood the safety factor. Some additional emphasis is needed here to explain why they are "overestimating" -- where Matt's "good enough" would not be good enough to solve the problem.

Finally, some of those who did not remember that the kids added another seal to the equation were comfortable telling us that 8 (eight) x 6 = 54. Apparently, the higher number multiplication tables are not familiar enough to some of the students. If the point of the story is "estimation," then maybe smaller numbers should be used to establish the point more successfully.