DragonflyTV
Formative Evaluation
Of Three Shows

Report for
Twin Cities Public Television

by
Barbara N. Flagg
Multimedia Research
33 Browns Lane
Bellport, NY 11713

with assistance of
Alice Bernard
Sandra Crow
Valerie Knight-Williams

Research Report No. 02-003
April 17, 2002

TABLE OF CONTENTS

EXECUTIVE SUMMARY OF FORMATIVE EVALUATION OF "DRAGONFLYTV" SERIES

MULTIMEDIA RESEARCH • APRIL, 2002

With support from the National Science Foundation and Best Buy, Twin Cities Public Television has produced a science series for children. Entitled "DragonflyTV," the series presents real kids engaged in real science to promote the process of science inquiry in viewers. This formative evaluation gathered feedback from 5th and 6th graders in response to three 28 minute videos, presented one per week. The following were the general research goals:

- To evaluate change in student ability to design an experiment;
- To measure change in interest in doing one's own science investigation;
- To determine differences in student ratings of the importance of four procedures in science investigation;
- To look for impact on attitude toward the corporate sponsor's involvement in helping kids learn;
- To assess appeal of the series of three shows overall and of individual segments.

Sample and procedure. Almost 300 5th and 6th grade students participated, drawn from Springfield, PA; Miami Beach, FL; and Elk Grove, CA. The sample comprised half females and two-fifths minorities. Students completed a pre-viewing survey; viewed one show per week for three weeks followed immediately by rating surveys, and then completed a post-series survey. A subsample of 48 randomly chosen students was interviewed individually both before and after the series to assess understanding of the design and implementation of experiments.

Results

Impact of Viewing DragonflyTV on Ability to Design an Experiment

Viewing *DragonflyTV* significantly improved students' ability to think about the design and implementation of an experiment. After seeing only three shows, over half (58%) of 48 interviewed students showed increased understanding of experimental design; understanding remained the same for 27% and decreased for 15%. After seeing the shows, students were significantly more likely to suggest comparing more than two items in their experiment; better able to explain a data collection procedure; more likely to suggest writing down or charting data; and better able to make interpretations. When asked if they learned anything new about how to design an experiment from watching the series, one-quarter did not think so, but 37% said they learned to use graphs, charts or tables and about 10% each noted the importance of writing things down; having an adult with you; going to different places to experiment and being well organized.

Impact of DragonflyTV on Interest in Doing Own Investigation

Viewing *DragonflyTV* significantly influenced children's interest in doing their own science investigation. Over one-third of students were more interested in doing a science experiment after viewing the *DragonflyTV* shows than they were before seeing the series. Minority viewers were

affected the most with 41% showing increased interest. At least 30% of viewers "very much" liked the idea of their own experiment after viewing three shows.

Impact of DragonflyTV on Perceived Importance of Four Procedures in Investigation

Viewing *DragonflyTV* significantly influenced children's rating of the importance of charting one's findings. One-third of non-viewers felt making a chart was "very important," whereas one-half of viewers thought so. Viewing did not have a significant impact on students' understanding of the importance of writing down what happens in an experiment, repeating an experiment or doing everything the same each time you do an experiment. The latter three strategies were less emphasized in the particular episodes viewed, whereas making charts was a significant aspect of the shows. "Writing down what happens" was already considered "very important" by 70% of respondents before seeing any episodes, and one-third of the non-viewers felt that way about the other research strategies.

DragonflyTV's Influence on Opinion of Corporate Sponsor, Best Buy

Viewing *DragonflyTV* positively influenced children's opinion of the involvement in education of the corporate sponsor, Best Buy. Children who watched *DragonflyTV* were three times more likely than non-viewers to agree that Best Buy is "very likely to care about helping kids learn new things." The average opinion of how much Best Buy "cares about helping kids learn" improved: On a scale of 1 to 4, where 4 meant "very likely," children who did *not* watch *DragonflyTV* ranked Best Buy at 2.4, the same rating as Target and Wal-Mart. But children who had watched three episodes of *DragonflyTV* gave Best Buy an average rating of 3.2, approximately 30% higher than non-viewers.

Appeal and Clarity for Shows 1, 2 and 3

DragonflyTV appealed to a significant majority of students. More than 80% of 5th graders and 70% of 6th graders gave positive ratings for ten major segments within the three episodes. Overall, 5th graders liked the major segments of *DragonflyTV* significantly more than 6th graders, and there were significant variations in the ratings for individual segments. Students rated Dogs and the Know How-Cutting Rock as the most appealing segments. Girls favored Otters and Forecasting Weather, whereas boys favored Tornado Model and Know How -Thief/fog. Both girls and boys were least interested in Know How-Airport Birds. Those who already watch science television rated segments higher than those who do not watch science television.

When asked what they liked or disliked about the series of shows, viewers most often mentioned particular segments because they liked or disliked the content (e.g., like dogs; dislike rocks).

Overall, viewers felt the content of the segments was clear to them, giving clarity ratings that were as high or higher than the appeal ratings. Generally, 5th graders and those who watch science television rated parts as clearer than 6th graders and those who do not watch science television.

INTRODUCTION

This formative evaluation gathered feedback from fifth and sixth graders in response to three 28-minute videos from the new children's science series, *DragonflyTV*. The general goals for the research were as follows:

1. To evaluate change in student ability to design an experiment;
2. To measure change in interest in doing one's own science investigation;
3. To determine differences in student ratings of the importance of four procedures in science investigation;
4. To look for impact of viewing on attitude toward the corporate underwriter's involvement in helping kids learn;
5. To assess appeal of the series of three shows overall and of individual segments.

METHOD

Episodes for viewing

The producers of the *DragonflyTV* series selected three episodes, designed to offer a range of subject matter. The themes of the episodes were "Animal Behavior," "Rocks and Minerals," and "Weather." They were episodes #3, #5 and #7 from the first season of 13 shows, but they were viewed in the order of #5, #3 and #7. The table below outlines the shows viewed, their order, content and type of research segments. "Major segments" were 5-8 minute investigations conducted by children, ages 10 to 13. In the first and second show viewed by students, the major segments, Kayaking and Otters, were presented in two parts. "Know How" segments were short, applied-technology riddles. These were posed at the beginning of the show, investigated by children mid-way through the show, and answered at the end. The riddle questions were "How can you cut rock without a saw?" "How do you keep birds off airport runways?" and "How do you stop a thief with weather?"[1] The scientist profiles present adult scientists who explain their job in their research setting.

Episode	Show order	Major segment	Know How	Major segment	Scientist Profile	Major segment
Rocks & Minerals	1	Kayaking (Experimental)	Cutting Rock	Rock climbing (Observational)	Paleontologist	Kayaking
Animal Behavior	2	Otters (Observational)	Airport Birds	Dog Behavior (Experimental)	Otter Aquarist	Otters
Weather	3	F-scale (Observational)	Weather stops thief	Tornado Model (Experimental)	Storm Chaser	Forecasting (Experimental)

[1] Answers: with a high-pressure water jet; with trained dogs or stuffed coyotes; with a household fog-generator.

Sample

Three sites were used for this study:

- The eastern site in Springfield, PA, is a 5th-8th grade school drawing from a middle to upper middle class suburb of Philadelphia. The 5th-6th grades are somewhat segregated from the older students and the school is 89% white.

- The southern site in Miami Beach, FL, is a pre-K-6th grade school drawing from very low to very high incomes in the Middle Beach area across from downtown Miami. The student body is two-thirds Hispanic.

- The western site in Elk Grove, CA, is a K-6th grade school drawing from a middle to lower middle class suburb of Sacramento. The 5th-6th grade student body is 60% white with 20% of students reporting mixed ethnic backgrounds and 9%, Asian.

Viewing samples for each show. Two fifth grade classes and two sixth grade classes were recruited to view the three programs at each of three sites: Sacramento, CA (n = 95); Miami Beach, FL (n = 100); and Springfield, PA (n = 102). All students who viewed and completed a show survey were included in the appeal analysis for that particular show (research goal #5 above). Students who completed both surveys for show 1 and show 3 and viewed all three shows were included in the analysis of pre-post research questions (research goals #2, 3, 4 & 5).

Nearly 300 students viewed each show, with about half female and two-fifths minorities including 20% Hispanics and the remainder African-American, Asian or mixed heritage. [2] The table below indicates sample distributions for individual shows.

Viewers	**Show 1 N = 295**	**Show 2 N = 281**	**Show 3 N = 284**
Fifth grade: Total #	140	132	133
% Females	44%	46%	46%
% Minorities	44%	39%	41%
% watch science TV[3]	70%	69%	70%
Sixth grade: Total #	155	149	151
% Females	53%	53%	53%
% Minorities	37%	37%	38%
% watch science TV	57%	57%	56%

[2] The National Center for Education Statistics reports in its most recent analysis in 1999 that 38% of students were part of a minority group with Hispanics making up about 16%. Our minority sample is slightly higher than the 1999 statistics - but since Hispanics were expected to increase their presence, the sample is probably on target for a 2002 distribution.

[3] This question was "do you ever watch any TV shows about science?" The 5th grade was significantly more likely than the 6th grade to report watching science television (70% of 5th grade vs. 57% of 6th grade watch science TV; Fisher Exact test p = .02).

Interviewed sample. Parental permission for a ten-minute individual pre- and post-viewing interview (to address research goal #1) was requested from all participating classes. Almost 100% of students returned signed slips. The interviewed sample was randomly chosen from the returned permission slips, stratified by gender and ethnicity. The goal for each site was to interview four 5th grade boys, four 5th grade girls, four 6th grade boys, and four 6th grade girls. The CA site provided African-American minority interviewees in each subgroup; the FL site provided Hispanic minority interviewees in each subgroup. A few extra pre-viewing interviews were carried out to cover the possibility of dropout or absenteeism. All interviewed students viewed the three shows and completed all questionnaires. As planned, the final sample of 48 students interviewed both before and after viewing three shows included equal numbers from each grade and equal gender groups within each grade. Minorities comprised 23% of the sample: 13% Hispanic; 6% African-American; and 4% mixed ethnicity.

Procedure

The procedure followed is outlined in the table and explained in the text that follows.

Week	5th grade -A	5th grade - B	6th grade - A	6th grade- B
1	Interview randomly chosen boys & girls with permission			
2	Pre-survey Sponsor View Show 1 Post-survey Show 1	Pre-survey Import View Show 1 Post-survey Show 1	Pre-survey Sponsor View Show 1 Post-survey Show 1	Pre-survey Import View Show 1 Post-survey Show 1
3	View Show 2 Post-survey Show 2 for all classes.			
4	View Show 3 Post-survey Show 3 Post-survey 3 Import Post-interview same kids	View Show 3 Post-survey Show 3 Post-survey Sponsor Post-interview same	View Show 3 Post-survey Show 3 Post-survey 3 Import Post-interview same	View Show 3 Post-survey Show 3 Post-survey Sponsor Post-interview same

During week one, 16+ students were interviewed about experimental design (see interview protocol and scoring in Appendix A). A random half of the interviewees were to design an experiment to find out why some balls bounce better than others, and the other half were to design an experiment to find out why some rabbits jump better than others. During week four, after viewing all three shows, the same students were interviewed with the opposite design question. The pre and post-viewing interviews with the same students were used to assess research goal #1.

During week two, all students completed a pre-viewing survey that asked:
- demographic questions;
- how much students would like to do a science experiment of their own;
- and asked one of two questions - Sponsorship (Sponsor) or Importance of Methodology (Import).
  - The Sponsorship question asked how much the stores Sears, Wal-Mart, Target and Best Buy care about helping kids like themselves learn new things.
  - The Import question asked how important it is to
    - (a) repeat the same experiment several times;
    - (b) write down whatever happens in your experiment;

- (c) do everything the same each time you do your experiment; and
- (d) put what you find out in a chart or a graph.

A random half of the sample, stratified by grade, answered the Sponsor question in the pre-viewing survey in week one and then answered the Import question in the post-viewing survey in week four, and vice versa. Additionally, the four store names and four experimental methods were presented in four different orders within the classes so as to counteract order effects. With the same sample, the pre and post questions were used to assess research goal #2, concerning interest in their own experiment. With different but equivalent samples, the pre and post questions were used to assess research goals # 3 about the importance of methodology and #4 about sponsorship.

During weeks two, three and four, after viewing each program, all students completed rating questions about the appeal and clarity of the specific segments of the shows (research goal #5). After viewing show 3, all students gave their written opinions about the full series.

Schools received an honorarium for participating in the evaluation and retained videos of the programs. It should be noted that the broadcast of *DragonflyTV* was occurring once weekly during the evaluation period, although no mention was made of this fact to students and no effort was made to determine whether students watched out of school or not.

Data Analysis

Pre and post-interviews about experimental design were transcribed verbatim by interviewers and scored independently by Dr. Flagg. Interviewers were not aware of the scoring criteria. The scoring criteria is more specifically described in Appendix A but focused on whether the child:

- Sets up a comparison of at least three items
- Describes one or more independent variables
- Describes one or more dependent variables
- Describes procedures of experiment, method of measuring dependent variable
- Defines more specifically how to measure dependent variable
- Suggests repetition of measurement or experiment
- Describes recording of data or method of saving information
- Discusses interpretation of data collected
- Describes possible problem and solution with experimental design

Means for research questions were compared with parametric tests (paired or two-sample t-tests, as appropriate). Pre-post comparisons were also made with non-parametric statistics[4] (one-tailed paired signed ranks test and McNemar test for significance of changes). All relationships were analyzed for statistical significance, which is reported if p values are less than .05. Variables explored included grade, gender, ethnicity and reported watching of science television. Keep in mind in reading the results that differences found between those who reported watching science television and those who did not report watching are closely related to grade differences, since significantly more 5th than 6th graders reported watching science television (70% vs. 57%).

[4] Parametric tests assume certain conditions about the population and require measurement at least in an interval scale. Non-parametric statistics are used when the assumptions of parametric tests may not be met and when data are in ordinal or nominal scales.

RESULTS: IMPACT OF VIEWING THREE SHOWS

Impact of viewing *DragonflyTV* on ability to design an experiment

Viewing *DragonflyTV* significantly improved students' ability to think about the design and implementation of an experiment. The average ability improved from 3.7 (out of a possible 11) before viewing to 4.9 after viewing, an increase of 32%. After seeing only three shows, over half (58%) of 48 interviewed students showed increased understanding of experimental design; understanding remained the same for 37% and decreased for 15%. After seeing the shows, students were significantly more likely to suggest comparing more than two items in their experiment; better able to explain a data collection procedure; more likely to suggest writing down or charting data; and better able to make interpretations. When asked if they learned anything new about how to design an experiment from watching the series, one-quarter did not think so, but 37% said they learned to use graphs, charts or tables and about 10% each noted the importance of writing things down; having an adult with you; going to different places to experiment and being well organized.

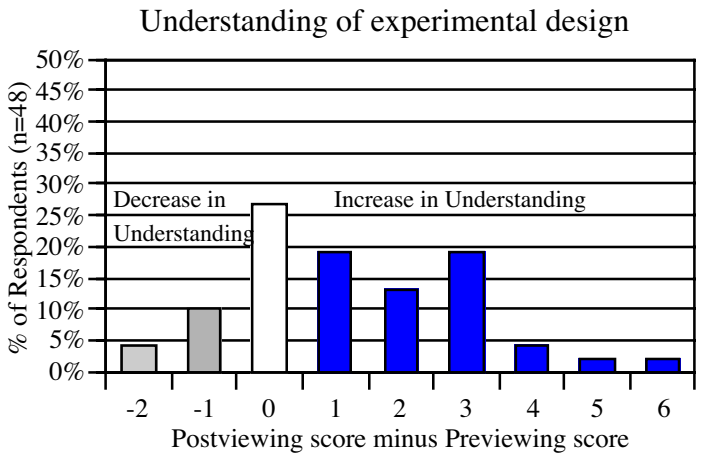
A randomly chosen subset of 48 viewers was interviewed before and after seeing the three shows. This subset included four 5th grade boys, four 5th grade girls, four 6th grade boys and four 6th grade girls at each of the three sites. Two equivalent open-ended interviews were developed. One scenario asked students to design an experiment to find out why some balls bounce better than others, and the other scenario asked them to design an experiment to find out why some rabbits jump better than others. To elicit students' own spontaneous ideas, very few prompts were (see Appendix A for interview specifics). At each site, a random two students in each grade/gender category were interviewed with "Balls" as a pre-test, and a random two students in each grade/gender category were interviewed with "Rabbits" as a pre-test. After the three shows were viewed, the same students were interviewed with the opposite scenario. This procedure was followed for each grade/gender subgroup so that each subgroup was balanced for scenario type. Students were interviewed and tape-recorded individually in a quiet place for approximately 10 to 12 minutes.

The total of 48 pre and post-interviews were transcribed verbatim by interviewers and scored independently by a non-interviewing researcher. Interviewers were not aware of the scoring criteria, which is more specifically described in Appendix A but focused on the following:

- Sets up a comparison of at least three items = 1 point
- Describes one or more independent variables = 1 or 2 points
- Describes one or more dependent variables = 1 or 2 points
- Describes procedures of experiment, method of measuring dependent variable = 1 pt
- Defines more specifically how to measure dependent variable = 1 point
- Suggests repetition of measurement or experiment = 1 point
- Describes recording of data or method of saving information = 1 point
- Discusses interpretation of data collected = 1 point
- Describes possible problem and solution with experimental design = 1 point

An interview score could range from 0 to 11. The pre-viewing scores of 48 students ranged from 0 to 8; the mean score was 3.7. There was no significant difference between the mean scores of the two different scenarios before viewing *DragonflyTV*. The post-viewing scores also ranged 0 to 8; the mean score after seeing the three shows was 4.9. There was no significant difference between the mean scores of the two scenarios after viewing *DragonflyTV*. The post-viewing mean was significantly higher than the pre-viewing mean score,[5] indicating that, on average, students who viewed *DragonflyTV* increased their understanding of how to design and carry out an experiment.

Out of 48 interviewed students, post-viewing understanding was greater than pre-viewing understanding in 58% of the sample (see dark bars in chart to the right). Understanding did not change for 27% of students (see white bar in chart to the right). Understanding decreased for 15% (see gray bars to the left side of chart). Viewing *DragonflyTV* significantly increased student understanding of how to design and carry out an experiment.[6]



A few examples of pre-viewing and post-viewing interviews are provided below with annotation about scoring:

Pre-interview of 6th grade boy for rabbit scenario:

Researcher: I'd like you to help me investigate a problem. There are no right or wrong answers to this problem. I just want to hear your ideas and what you think I could do to investigate the problem. You're familiar with different kinds of jumping animals, right? [pause] Some animals jump better than others. Think about rabbits. I want to figure out why some rabbits jump better than others. Describe what kind of experiment you might set up to find out why some rabbits jump better than others.
Student: You can take two rabbits with one being more muscular than the other and see which one jumps better by like holding something up and seeing which one jumps better over it, like jumps higher.
Researcher: Okay, so imagine you have two rabbits with one being more muscular than the other, walk me through step by step what you would do to figure out why some rabbits jump better than others.
Student: Somehow you would set up something that would allow you to get them to jump over and then just do like I said.
Researcher: Can you tell me anything else you would do to help you come to a final conclusion for why some rabbits jump better than others?
Student: I would just study why some jump higher than others.
Researcher: Can you tell me anything more about that?
Student: No.
Researcher: Can you think of any problems doing the experiment the way that you did?
Student: It could be hard to get them to jump. And it could be hard to find the rabbits too maybe.

[Independent variable = musculature; Dependent variable = height of jump; Procedure = holds something up to jump over. Pre score = 3]

[5] Paired t - test, $p < .002$.

[6] One-tailed Wilcoxon matched-pairs signed ranks test looks at the direction and relative magnitude of the pre-post differences in ratings for individual students; $p = .0003$).

Post-interview of same 6th grade boy for balls scenario:

Researcher: I'd like you to help me investigate a problem like the ones you saw on the *DragonflyTV* programs. There are no right or wrong answers to this problem. I just want to hear your ideas and what you think I could do to investigate the problem. You're familiar with different kinds of bouncing balls, right? [pause] Some of them bounce better than others. I want to figure out why some balls bounce better than others. Describe what kind of experiment you might set up to find out why some balls bounce better than others.
Student: You can make a chart of each ball. You can write something down after you drop them from somewhere and see which one bounces better, the highest, and then you can write that down on a chart and compare them.
Researcher: Could you give me a little more detail about that?
Student: I'm not sure what I'd add.
Researcher: I want you to imagine that you have these different balls. Now just walk me through what you would do, step by step, to figure out why the balls bounce differently, some better than others.
Student: I'll get some of all sizes and see what bounces, and then I'd write them down on the chart and then I'll find something that I can drop them off of to see which one bounces off of and maybe measure it.
Researcher: Any more detail you can add about that?
Student: First of all I'd have to build something to drop them off of or just find something to drop them off of. Then you'll have to get the balls from somewhere and then just drop them. Then see how high they bounce and have someone measure it below you and then write it down on a chart and compare to see which one bounces better. Also, you can test like how much air each one has. Put an equal amount of air in each ball, the basketball, the bouncy ball, and then just drop them down and see which one bounces higher with the same amount of air.
Researcher: Can you tell me anything else that you would do to help you come to a final conclusion for why some balls bounce better than others?
Student: Maybe you can try it with a different basketball: Say the basketball bounces the highest, and see if it's like the certain kind of rubber it's made with or something that makes it bounce higher and then you can check to see if that does it or it's just the type of ball or size.
Researcher: Can you think of any problems doing the experiment the way that you did?
Student: You have to figure out how to get the balls and put the same amount of air in them.
Researcher: Any other problems?
Student: No.

[More than 2 balls; Independent variables = size and air; Dependent variable = height of bounce; Procedure = drop off of something he builds, have someone measure height of bounce below him; Record = write down and make chart. Post score = 6]

Pre-interview for 5th grade girl for ball scenario:

Researcher: I'd like you to help me investigate a problem. There are no right or wrong answers to this problem. I just want to hear your ideas and what you think I could do to investigate the problem. You're familiar with different kinds of bouncing balls, right? [pause] Some of them bounce better than others. I want to figure out why some balls bounce better than others. Describe what kind of experiment you might set up to find out why some balls bounce better than others.
Student: Maybe the ball might be like, one might be like older than the other, or maybe like newer than the other, and so the other one's worn out, or one may have more air than the other one and you can just test by taking some air out one ball and put some more air in another and see what the differences is. And then you can do it that way or you can take, try to find out about one ball by asking a person how old it is and then you can get another one that's a little newer and then bounce them and see how high they go.
Researcher: So it sounds like you're describing two experiments: one with old and new balls, and one with balls having various amounts of air in them?
Student: Yeah.
Researcher: Okay, let's take either of the two experiments. I want you to imagine that you have old and new balls, or balls with a little air and a lot of air. Now walk me through what you would do, step by step, to figure out why the balls bounce differently, some better than others.

Student: Well, of course, you would bounce them first and see which one would go higher. And then we would, we would be able to say that the one with less air was the lowest bouncing, and the one with more air bounced a little bit more better.
Researcher: Could you explain in more detail what you would do to figure out why balls with different amounts of air bounce in different ways?
Student: You could do that and you would, like, first see how they bounce and then maybe put more air in them and see then.
Researcher: Okay. Can you tell me anything else that you would do to help you come to a final conclusion for why some balls bounce better than others?
Student: Maybe, you might find out that we should make sure that there's a lot more air and then it can bounce higher than the one that you take just a little bit more air out of that don't bounce so high.
Researcher: Can you think of any problems doing the experiment this way?
Student: Maybe if when somewhere that you can, once you take too much air out that it wouldn't bounce at all, and when you put too much air in, it would pop.

[Independent variable = air; Dependent variable = height of bounce; Procedure = bounce them, see which goes higher, put more air in, bounce again. Pre score = 3]

Post-interview for same 5th grade girl for rabbit scenario

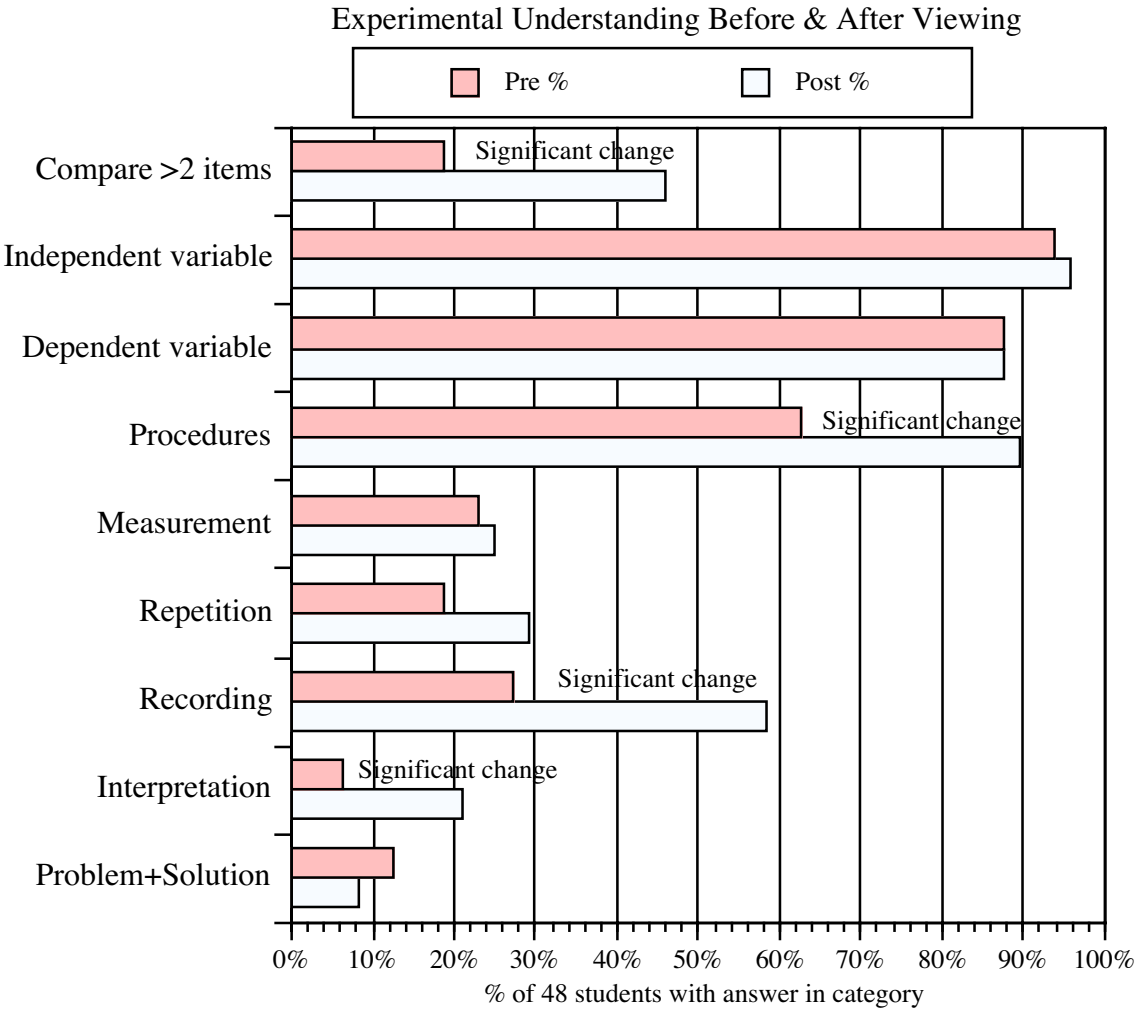
Researcher: I'd like you to help me investigate a problem like the ones you saw on the *DragonflyTV* programs. There are no right or wrong answers to this problem. I just want to hear your ideas and what you think I could do to investigate the problem. You're familiar with different kinds of jumping animals, right? [pause] Some animals jump better than others. Think about rabbits. I want to figure out why some rabbits jump better than others. Describe what kind of experiment you might set up to find out why some rabbits jump better than others.
Student: Um, I would take a little baby rabbit and then a really old rabbit and compare their jumping ability. The mommy and daddy rabbits of these rabbits would have to be the same kind so that I'm not taking into account the abilities of different kinds of rabbits.
Researcher: Okay, so I want you to imagine that you have a baby rabbit and an old rabbit of the same kind of rabbit. Now walk me through what you would do, step by step, to figure out why some rabbits jump better than others.
Student: I would take the baby one and the grown-up one and somehow I would need to find a way to get them to jump and then I would measure which one jumped the highest and which one jumped the longest, cuz that might be different depending on how old a rabbit is. I would probably watch them for a couple of hours and then I'd like try to estimate how high they are jumping and then I would probably draw a graph with different colors to show me how the different ones are jumping.
Researcher: Okay. Can you tell me anything else that you would do to help you come to a final conclusion for why some rabbits jump better than others?
Student: I would probably go to a zoo or places that have rabbits and I would talk to them about all the different things you should know about rabbit jumping. And I might go to a library to read about different kinds of rabbits so I can make more experiment even more, even more like accurate.
Researcher: Okay, can you think of any problems doing the experiment the way you did?
Student: Well, I may be missing something about rabbits or about the different kinds of rabbits and what makes some rabbits jump better. I may need to read or do some research first before I do my experiment. And there may be something about the needs of little ones and big ones that are important to jumping, like how much they eat or when they sleep and when they like to jump. I'm just not sure I have a really good question to research yet and I may need to find out more about rabbits to come up with something better.

[Independent variable = age; Dependent variables = height and length of jump; Procedure = find way to get them to jump, watch for hours, estimate how high jumping. Record = draw a graph with different colors. Post score = 5 [7]]

[7] Note that background reading and an observational study of sorts is proposed, which are both activities that Dragonfly investigators did. The scoring did not look at such issues.

The chart below presents the percentages of interviewed students who were able to give an acceptable answer for each of the nine scoring categories before and after viewing *DragonflyTV*. Prior to viewing, 94% of students could think of some sort of independent variable, and 88% could describe a dependent variable; these percentages did not change significantly after seeing the series. For the interviewed sample as a whole, four categories of experimental design did improve significantly after viewing *DragonflyTV*:[8]

- the suggestion to compare more than two items in their experiment;
- the ability to explain a data collection procedure;
- the desire to write down or chart data; and
- the ability to make interpretations, for a small but significant portion of the sample.



The significant change in the interviewees' suggestions to record their data reflects another finding in this study, described later in the report: One-half of students felt making a chart was "very important" after seeing the three programs, whereas only one-third thought so before seeing the shows. One student when asked if she learned anything new about how to design an experiment from watching the shows suggested directly that she learned about charts, a method she had not thought about in her pre-interview:

> "They talked about graphs. That's where I got the idea [to use a graph with different colors in rabbit post-interview] because when we did the ball experiment [in pre-interview], I never thought about that, but

[8] McNemar tests for significance of changes, all $p < .05$.

then the show did something about using graphs, and I thought that was a good idea to write your findings down on paper so you won't forget."

As mentioned previously, 15% of the interviewees (7 students) decreased in their interview score. There were no strong similarities among these students - the group included both genders, both grades, all states, minority as well as white students, and those who were presented with the ball scenario first and the rabbit scenario first. The decrease in pre to post scores occurred for both high scoring students (8 to 7) and low scoring students (3 to 2). The only seemingly consistent finding was that 4 of the 7 students failed to mention a dependent variable in their post-viewing experimental design. There is nothing to indicate any one factor related to the decrease in scores.

At the end of the post-interview, students were asked if they learned anything new about how to design an experiment from watching the *DragonflyTV* shows, and if so, what did they learn. One-quarter of those asked did not think they had learned anything new. Of those who did report learning,

- 37% learned to use graphs, charts or tables:
  "I wouldn't have thought about using graphs in outdoor projects because I usually think about them in a science lab."
  "It's a good idea to do charts because it helps you to organize all your thoughts and your procedures."
- 11% noted that the kids wrote down what they were doing and their data;
- 9% suggested that having an adult is a good idea;
- 9% learned you can go to lots of different places to experiment:
  "An experiment is not just something that you do in a house at a desk, you have to go all over and examine stuff."
  "It showed a lot of different places. I wouldn't have thought about doing experiments in those kinds of places."
- 9% observed that you have to "have it all organized."
- 7% described trying things out more than once;
- 7% warned to be "careful and neat;"
- 4% reflected that there are many different ways to do experiments, "there isn't just one way that is right."
- 4% noted that some experiments might not work but "that that's okay."

Then there were a few unique responses revealing interesting unplanned effects:

5th grade boy, with 8 on pre and post, enjoyed thinking about better experiments:
"The dogs one was kind of hard to do because they didn't eliminate as many variables as possible. They could have eliminated more, but I guess they just wanted to look at their dogs, so that was okay. I just liked thinking through what the kids may have missed and how to do the experiment better. It's a lot better than just sitting in class and taking tests and doing book stuff."

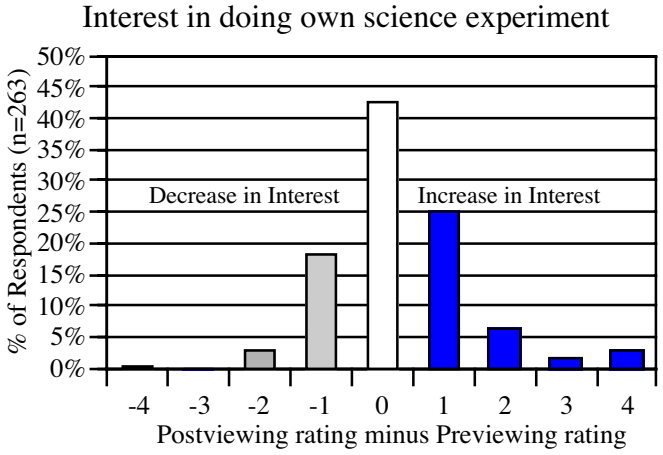
5th grade girl, with 4 on pre and post, observed that projects don't need a display:
"What I really learned was like how to present it, liked project your voice and everything, because some projects, like the little tornado, they sort of showed how they did it step by step. They sort of explained it to you really, really really clear and it got in my head. I would think a science project -- oh, boards, I need paste, I need glue -- here instead you just like show it to everyone in front of the class instead - you can just do it without a board."

5th grade boy, with 2 on pre and 3 on post, liked the series' data interpretation:
"They showed me how they got to the conclusion. Lots of shows say we got to the conclusion, but they don't show me specifically how."

Impact of *DragonflyTV* on interest in doing own investigation

Viewing *DragonflyTV* significantly influenced children's interest in doing their own science investigation. Over one-third of students were more interested in doing a science experiment after viewing the *DragonflyTV* shows than they were before seeing the series. Minority viewers were affected the most with 41% showing increased interest. At least 30% of viewers "very much" liked the idea of their own experiment after viewing three shows.

The same students were asked both before and after viewing the series of three shows how much they would like to do a science experiment of their own. A total sample of 263 children completed both pre and post questions and viewed all three shows. Post-viewing interest was greater than pre-viewing interest for 36% of the sample (see dark bars in chart to the right). Interest in doing their own experiment did not change after viewing for 43% of the sample (see white bar in chart). Interest decreased for 21% (see gray bars on left side of chart). Viewing *DragonflyTV* significantly increased student interest in doing an experiment.[9]



The largest impact was on minority viewers: 41% were more interested in doing their own science experiment after viewing the three shows. Indeed, in the post-interviewing procedure, a 5th grade Hispanic girl offered that she had tried to make her own tornado model:

"I tried to do it, but I didn't have enough fans, so I couldn't finish it."

Also, a 5th grade Hispanic boy who was interviewed reported interest in trying to grow crystals:

"And I know the one that said, Did you know that the salt in the water, I was going to try that on Saturday."

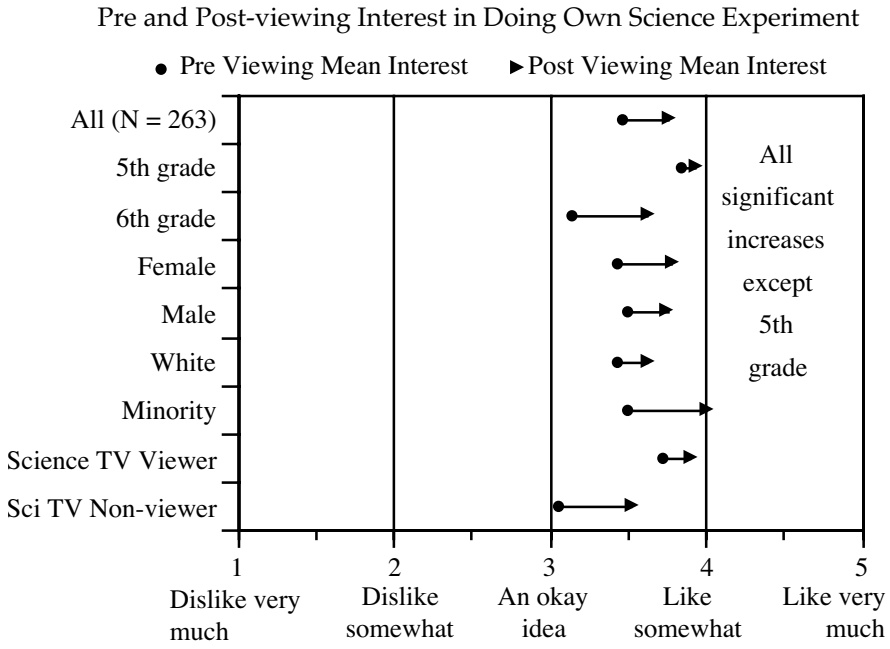
Additionally, a white 6th grade boy in his interview described replicating the dog activity:

"Me and my friend. We were at my friend's house, and he has a golden retriever. Wasn't it a golden retriever on the video? And we did an experiment. We got this bone and put it under a cover, and he got it right off. And then we put the cover on top of the dog, and he was walking around for like an hour. And then we put it like on a harder surface, and he couldn't like get it out. It was pretty fun. To see how smart the dog was."

So, three of our 48 interviewed students spontaneously reported being prompted by their viewing of *DragonflyTV* to model the investigative activities that they saw. Possibly more students were so affected but the procedure did not ask directly.

[9] One-tailed Wilcoxon matched-pairs signed ranks test looks at the direction and relative magnitude of the pre-post differences in ratings for individual students; $\underline{p} = .0009$).

The same rating data may be analyzed in a different manner by looking at average interest ratings before viewing and after viewing. The chart below indicates pre-viewing interest means with a circle • and post-viewing means with a triangle ▸ for all subgroups. Viewing *DragonflyTV* significantly increased mean interest in doing an investigation for the sample as a whole and for all subgroups, except the 5th grade, which already had high interest prior to viewing the series. Thus, 6th graders; females and males; whites and minorities; science TV viewers and non-viewers; were on average more interested in doing their own experiment after seeing the three shows.



When we examine the most positive end of the rating scale, we find that before seeing *DragonflyTV*, 19% of the whole sample liked the idea of doing their own experiment "very much (5)," whereas after viewing the shows, 30% liked the idea "very much (5)."

Impact of *DragonflyTV* on perceived importance of four procedures in investigation

Viewing *DragonflyTV* significantly influenced children's rating of the importance of charting one's findings. One-third of non-viewers felt making a chart was "very important," whereas one-half of viewers thought so. Viewing did not have a significant impact on students' understanding of the importance of writing down what happens in an experiment, repeating an experiment or doing everything the same each time you do an experiment. The latter three strategies were less emphasized in the episodes viewed, whereas making a chart was a significant aspect of the shows.

Students were given a list of four methods of scientific investigation and asked "how important each method is to doing a good experiment," as shown in the question below:

There are a lot of different methods to do a good experiment. Some of the methods are described below. Circle one number that tells how important each method is to doing a good experiment.

	Very important	Somewhat important	Somewhat unimportant	Very unimportant	Don't know
How important is it to repeat the same experiment several times?	4	3	2	1	0
How important is it to write down whatever happens in your experiment?	4	3	2	1	0
How important is it to do everything the same each time you do your experiment?	4	3	2	1	0
How important is it to put what you find out in a chart or a graph?	4	3	2	1	0

To control for a cueing effect, half of the classes in each grade were chosen randomly to answer the question before viewing (non-viewers) and the other half of classes answered the question after viewing (viewers). To control for primacy effects, methods were ordered in four different ways and the orders were presented randomly to students in each class. The analysis sample comprised those who had viewed all three programs. The pre and post mean responses for each method are listed in the table below:

Mean ratings of perceived importance of experimental method

Method	**Non-viewers** (n = 122)	**Viewers** (n = 147)	
How important is it to put what you find out in a chart or a graph?	3.0	3.3	**Significant diff.** *t*-test, $p \leq .004$
How important is it to write down whatever happens in your experiment?	3.5	3.7	
How important is it to do everything the same each time you do your experiment?	2.9	2.9	
How important is it to repeat the same experiment several times?	2.8	3.0	

In the previous table, all pre-viewing mean ratings are fairly high (2.8 to 3.0 out of 4). Only the means for using a chart or graph show a significant difference between non-viewers and viewers. We can conclude that viewing *DragonflyTV* helped children place a higher importance on the process of charting one's findings.

The table below presents the percentages of students who rated the procedures as "very important." The first row indicates that writing down what happens was already considered very important by 70% of respondents before seeing any shows. About one-third of the sample felt that way about the other three methods. The only significant difference is between the 36% of non-viewers who thought using a chart was "very important" and the 51% of viewers who thought a chart was "very important."

Percent of students who rated procedure as "very important"

	Non-Viewers	Viewers
How important is it to write down whatever happens in your experiment?	70%	73%
How important is it to do everything the same each time you do your experiment?	39%	39%
How important is it to put what you find out in a chart or a graph?	36%	51%
How important is it to repeat the same experiment several times?	27%	35%

In the production of the series, charting was stressed consistently across all investigations but the other three investigative elements were not so consistently stressed. What might be inferred from these findings is that students understand the strategies presented in the video in proportion to how much stress is placed on a given strategy.

DragonflyTV's influence on opinion of corporate sponsor, Best Buy

Viewing *DragonflyTV* positively influenced children's opinion of the corporate sponsor's involvement in education. Children who watched *DragonflyTV* were three times more likely than non-viewers to agree that Best Buy is "very likely to care about helping kids learn new things." The average opinion of how much Best Buy "cares about helping kids learn" improved: On a scale of 1 to 4, where 4 meant "very likely," children who did *not* watch *DragonflyTV* ranked Best Buy at 2.4, the same rating as Target and Wal-Mart. But children who had watched three episodes of *DragonflyTV* gave Best Buy an average rating of 3.2, approximately 30% higher than non-viewers.

Before and after viewing the three shows, students were given a list of four stores and asked "how likely it is that each of the stores cares about helping kids like themselves learn new things," as shown in the question below:

There are stores that want to help kids learn. Some stores are listed below. Circle one number that tells how likely it is that each of the stores cares about helping kids like you learn new things.

	Very likely	Somewhat likely	Somewhat Unlikely	Very unlikely	Never heard of store
How likely is it that Best Buy cares about helping kids learn new things?	4	3	2	1	0
How likely is it that Sears cares about helping kids learn new things?	4	3	2	1	0
How likely is it that Wal-Mart cares about helping kids learn new things?	4	3	2	1	0
How likely is it that Target cares about helping kids learn new things?	4	3	2	1	0

To control for primacy effects, store names were ordered in four different ways and the orders were presented randomly to students in each class. To control for a cueing effect, half of the classes in each grade were chosen randomly to answer the question before viewing (non-viewers) and the other half of classes answered the question after viewing (viewers). The analysis sample comprised those who had "heard of" all four stores and who had viewed all three programs.

The pre and post mean responses for each store are listed in the table below:

Mean ratings of likelihood that stores care about helping kids learn new things

Store	**Non-viewers** (n = 140)	**Viewers** (n = 119)	
Best Buy	2.4	3.2	**Significant diff.** *t*-test, $p \leq .0001$
Sears	2.0	2.2	
Wal-Mart	2.5	2.6	
Target	2.6	2.7	

Only the Best Buy mean ratings show a significant difference between non-viewers and viewers. Also, the Best Buy post-viewing mean was significantly higher than the post-viewing means of the other three stores. Students recognized after viewing only three shows that Best Buy cares about helping kids learn new things.[10]

The viewers of *DragonflyTV* were three times more likely than non-viewers to agree that Best Buy is "very likely to care about helping kids learn new things." Of the pre-viewing group, 16% chose the "very likely" response. After three viewings of the series, 48% chose the "very likely" response for Best Buy. Clearly, viewing *DragonflyTV* positively influenced children's opinion of Best Buy's support of education.

[10] The pre-post mean differences for Best Buy hold true for all sub-group comparisons within grade, gender, ethnicity and state, except for the Florida sample, which had a high opinion of Best Buy prior to viewing (FL pre mean = 2.8; post mean = 3.1).

RESULTS: APPEAL AND CLARITY FOR SHOWS 1, 2, 3

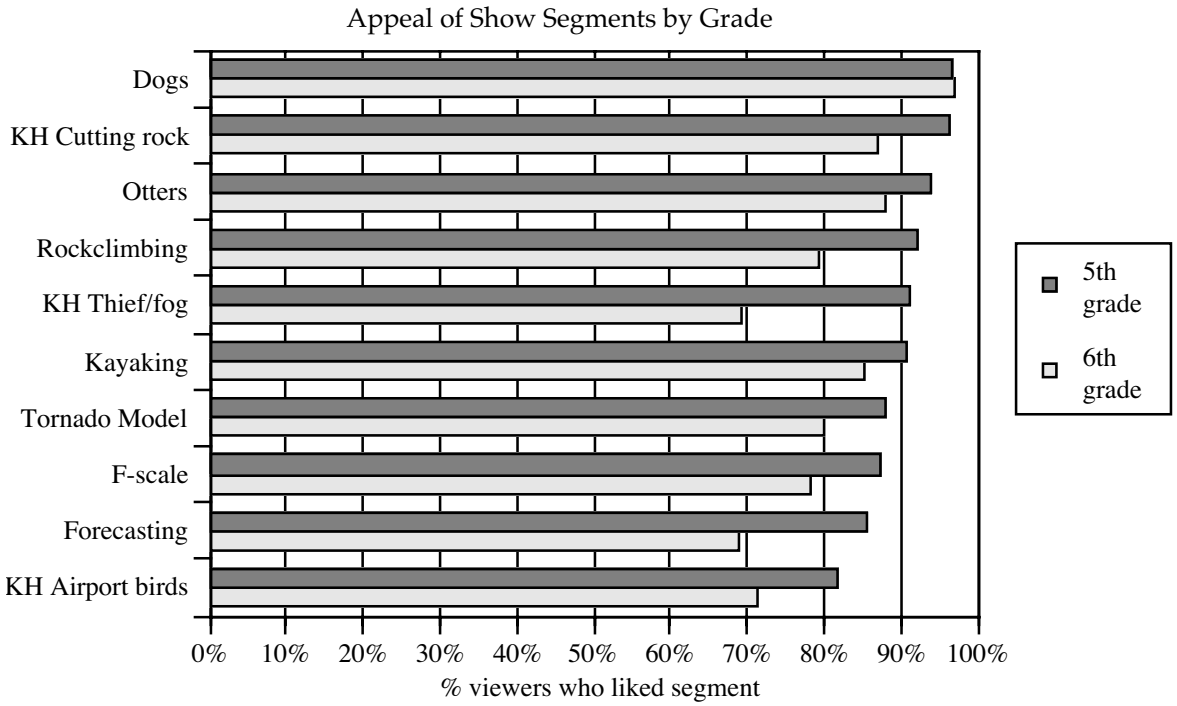
DragonflyTV appealed to a significant majority of students. More than 80% of 5th graders and 70% of 6th graders gave positive ratings for ten major segments within the three episodes. Overall, 5th graders liked the major segments of *DragonflyTV* significantly more than 6th graders, and there were significant variations in the ratings for individual segments. Students rated Dogs and the Know How-Cutting Rock as the most appealing segments. Girls favored Otters and Forecasting Weather, whereas boys favored Tornado Model and Know How - Thief/fog. Both girls and boys were least interested in Know How-Airport Birds. Those who already watch science television rated segments higher than those who do not watch science television (this variable was closely related to grade differences since 70% of 5th graders and only 57% of 6th graders reported watching science television).

When asked what they liked or disliked about the series of shows, viewers most often mentioned particular segments because they liked or disliked the content (e.g., like dogs; dislike rocks).

Overall, viewers felt the content of the segments was clear to them, giving clarity ratings that were as high or higher than the appeal ratings. Producers reported that project advisors were concerned that the video style and use of pop music might impair student understanding, but this appears not to be the case. Generally, 5th graders and those who watch science TV rated parts as clearer than 6th graders and those who do not watch science TV.

Appeal ratings of show segments

After viewing each of the three shows, viewers rated each major part in the show on a five-point appeal scale. The study did not ask for an overall appeal rating of the show, rather the producers requested a more detailed response to individual segments. The chart that follows presents the percent of 5th and 6th graders who were positive in their appeal ratings, saying "liked a lot;" "liked somewhat;" or "it's ok." The segments are ordered by 5th grade percentages, which range from 97% to 82%. The 6th grade range is 97% to 71%. More than 80% of 5th graders and 70% of 6th graders rated 10 of the 10 segments positively.


Appeal of Show Segments by Grade
Dogs
KH Cutting rock
Otters
Rockclimbing
KH Thief/fog
Kayaking
Tornado Model
F-scale
Forecasting
KH Airport birds
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
% viewers who liked segment
5th grade
6th grade

Average appeal ratings for the series' main segments ranged from a 4.4 to 3.2 out of a possible 5, as indicated in the second column in following table. The remaining columns indicate significant appeal differences for grade, gender, reported watching of science television, and ethnicity. To interpret the table, for example, $5^{th} > 6^{th}$ means that 5^{th} grade boys liked a segment significantly more than 6^{th} grade boys, and yes > no means that those who reported that they watch science television liked a segment significantly more than those who reported that they do not watch science television. [Specific means by subgroup are provided in tables in Appendix B.]

Fifth graders liked the segments significantly more than sixth graders, and those who reported watching science television also favored the segments more. Females liked three segments significantly more than males: Otters; Dogs; Forecasting weather. Fifth grade males liked three segments significantly more than 5^{th} grade females: KH-airport birds; KH-thief/fog; Tornado model. Ethnic differences appeared for only two segments for the fifth grade only: Hispanics liked KH-Cutting rock more than white students; Hispanics and African-Americans liked Rock Climbing more than white students.

Significant differences in mean appeal ratings for segments

Show 1 **Rocks and Minerals** (n = 295)	**Mean Appeal**	**Grade**	**Gender**	**Watch science TV or not**[11]	**Ethnicity**
Kayaking - why boulders are in rapids and not calmer waters	3.5	$5^{th} > 6^{th}$		Yes > No	
Know how - how do you cut through a rock without a saw	4.0	$5^{th} > 6^{th}$		Yes > No	For 5^{th} only Hisp > wh
Rock climbing - how do 3 types of rock affect climbing	3.5	$5^{th} > 6^{th}$		For 6^{th} only Yes > No	For 5^{th} only Hisp, Blk > wh
Show 2 **Animal Behavior** (n = 281)					
Otters - do otters groom more in the aquarium or in the wild	3.6	$5^{th} > 6^{th}$	F > M	Yes > No	
Know how - how do you keep birds off an airport runway	3.2	For M only $5^{th} > 6^{th}$	For 5^{th} only M > F	Yes > No	
Dogs - which dogs solve problems more quickly	4.4	$5^{th} > 6^{th}$	F > M	Yes > No	
Show 3 **Weather** (n = 284)					
F-Scale - how strong was the tornado in the town of Siren	3.5	$5^{th} > 6^{th}$		Yes > No	
Know how - how can weather stop a thief in their tracks	3.6	$5^{th} > 6^{th}$	For 5^{th} only M > F	Yes > No	
Tornado model - which wind is more important - top or side	3.7	For M only $5^{th} > 6^{th}$	For 5^{th} only M > F	Yes > No	
Forecasting weather - predict weather from nature clues	3.4	$5^{th} > 6^{th}$	F > M		

[11] Keep in mind that differences found between those who reported watching science television and those who did not report watching are closely related to grade differences, since significantly more 5^{th} than 6^{th} graders reported watching science television (70% vs. 57%).

The table below orders all ten show segments by their mean appeal ratings by subgroups of grade and gender, with the more appealing segments at the top. [Note that rows across do not indicate equivalent appeal; as an example, the segments with a 3.7 mean appeal are indicated in parentheses in the table.] All student subgroups rated Dogs and the Know How-Cutting Rock as the most appealing segments. The cells in lightest gray shading indicate that girls from both grades liked most the same four segments and liked least the same segment. The cells in darker gray shading indicate that boys from both grades liked most the same four segments and liked least the same two segments. Gender appears stronger in affecting order of segment preference than grade, although 5th graders overall liked the series and all individual segments more than 6th graders.

5th grade girls	6th grade girls	5th grade boys	6th grade boys
Dogs	Dogs	Dogs	Dogs
KH Cutting rock	KH Cutting rock (3.7)	KH Cutting rock	KH Cutting rock (3.7)
Otters	Otters	KH Thief/fog	Tornado model
Forecasting	Forecasting	Tornado model	KH Thief/fog
Rock climbing	Tornado model	Kayaking	F-Scale
KH Thief/fog (3.7)	Rock Climbing	F-Scale	Rock Climbing
Tornado model	Kayaking	Rock Climbing	Otters
Kayaking	F-Scale	Otters (3.7)	Kayaking
F Scale	KH Thief/fog	KH Airport birds	KH Airport birds
KH Airport birds	KH Airport birds	Forecasting	Forecasting

When asked what they liked or did not like about the series overall, most students mentioned a particular segment or show and said they liked or did not like the content; e.g., "liked the dog experiment because I love dogs;" or "disliked rock one because I don't like rocks or what kind they are." Responses traceable to a particular show or segment were categorized by show. The tables below present the percentages of each subgroup who commented positively and negatively about a show:

Distribution of positive comments about individual shows

Show	5th Female	6th Female	5th Male	6th Male	All
Rocks & Minerals	23%	11%	35%	13%	20%
Animal Behavior	56%	29%	33%	18%	33%
Weather	46%	40%	65%	29%	48%

Distribution of negative comments about individual shows

Show	5th Female	6th Female	5th Male	6th Male	All
Rocks & Minerals	23%	24%	10%	14%	17%
Animal Behavior	10%	8%	3%	7%	8%
Weather	31%	26%	28%	27%	28%

- Among the three shows, the third show viewed, Weather, received almost half of the positive comments but also most of the negative comments. Illustrative examples are quoted below.

  - Know How- thief/fog was commented about mostly by boys, particularly 5th grade (26% of 5th boys noted this segment). They felt the segment was "cool how the smoke came out;" "funny;" "awesome;" "interesting and made sense." As shown later in this report, this segment was highly ranked for clarity. On the other hand, 6th grade girls felt this segment was "dumb;" "stupid because you can't control weather;" "just plain impractical;" "farfetched;" "not detailed enough." These qualitative comments reflect the students' post-viewing quantitative ratings: 5th grade boys rated this segment third and 6th grade girls rated it ninth out of ten (see chart on previous page).

  - The Tornado model was described in positive terms most frequently by boys of both grades (15% of 6th grade boys noted this segment). They liked the project because it was "cool;" "interesting how they made it;" "a fun way to learn." A few liked the model making because it "tells me how it forms," and a few felt they could "do it in my own home." A few 5th grade females said they disliked this segment because the explanation was confusing to them.

  - Forecasting weather received positive comments from the girls, mostly about how they could do the project on their own (15% of 6th grade girls and 13% of 5th grade girls noted this segment):
    "because you can forecast by yourself, it was fun and I learned new things;"
    "something I would do sometime;"
    "funny when they were graphing the grandmother's toe and I found the curling hair interesting;"
    "most interesting, contained useful information;"
    "I understood what they were saying."
    In contrast, the boys (5th, 10%; 6th, 11%) and some girls (6th, 8%) described this segment as boring or confusing and that talking about "grandma's toe was weird" and "gross."

  - F-scale was enjoyed because viewers "love watching storms;" "learned how fast a tornado moves;" "could get what was going on." The few negative comments about this segment suggested that it was not clear enough; the overall clarity rating was on the low end of the ten segments.

- The second show viewed, Animal Behavior, received one-third of the positive comments and only 8% of the negative criticisms:
  - The Dogs segment was praised most frequently (30%) of all the ten parts and was favored highest across gender and grade subgroups. Students liked the segment because they "like dogs," because it was "interesting," and "fun" and "very clear." As shown later in this report, the Dogs segment received the highest mean clarity rating of all segments rated.

  - Otters was noted most by 6th grade girls (16%), who found the segment "interesting," or "cool" because they like animals. The few (2%) students who disliked the otter segment felt that it was "boring" and "a bit confusing."

  - Know How-Airport birds was described as "liked" by three boys only. Those who mentioned this segment in their negative commentary found it "boring" and "confusing;" it is ranked as one of the more confusing segments, as shown later.

- One-fifth of the sample described segments of the first show viewed (Rocks & Minerals) as ones they liked, and 17% described segments of this show that they disliked:
  - Kayaking received the most praise overall, but particularly from 5th grade boys (22%): "like kayaking;" "learned something;" "experiment was interesting and clear." About 6% of viewers wrote that this segment was "confusing," and as shown later in this report, Kayaking ranked quantitatively as least clear of the ten parts.

  - Rock climbing also was noted most positively by 5th grade boys (17%): "rocks are really cool and fun to climb;" "showed different kinds of rock and what better to climb on;" "because it talked about different kinds of rock and helped me in my science test;" "clear, and kids were the ones climbing, not grown-ups." About 6% of viewers said they were simply not interested in rocks or rock climbing.

  - Although Know How-Cutting Rock was rated quantitatively as a very appealing segment, only 5% of the sample commented on it in the open-ended question, noting that it was "amazing;" "cool;" "interesting;" and taught them something new. This segment did not elicit any negative criticism.

- Respondents also liked or disliked more general aspects of the series:
  - The music was quite popular with 6th graders (18%) who felt it was the kind of music that they listen to. A few students (4%) took exception to the theme song: "too happy and childish;" "weird;" "annoying;" "corny."

  - 16% of 6th graders and 4% of 5th graders described the jokes as "corny;" "stupid;" "dumb;" "weird;" "lame;" or "cheesy."

  - 13% of 6th graders and 8% of 5th graders were critical of the hosts, commenting on their script, their acting, and their attitude toward the audience. Sixth graders wrote:
    "they treated us as if we were little kids;" (3 students said similar thought)
    "they can get annoying;" (3)
    "hosts tried too hard to be funny;" (3)
    "hosts are weird;" (2)
    "hosts are corny;" (2)
    "hosts were really dumb and corny because of what they said and their actions;"
    "hosts said some cheesy stuff;"
    "how fake acting the two hosts were."

  - The experiments as a concept were praised mostly by 6th graders (11%), who wrote:
    "all the different activities because it's cool seeing them make charts and going to cool different places to find things;"
    "how you could make your own experiments;"
    "how they would teach you experiments;"
    "where they showed an experiment to do because I might do that on my spare time;"
    "how you investigate problems and solve them and the problems you choose are good too;"
    "shows had interesting experiments that I might try to do someday;"
    "cool experiments. I like to do stuff."

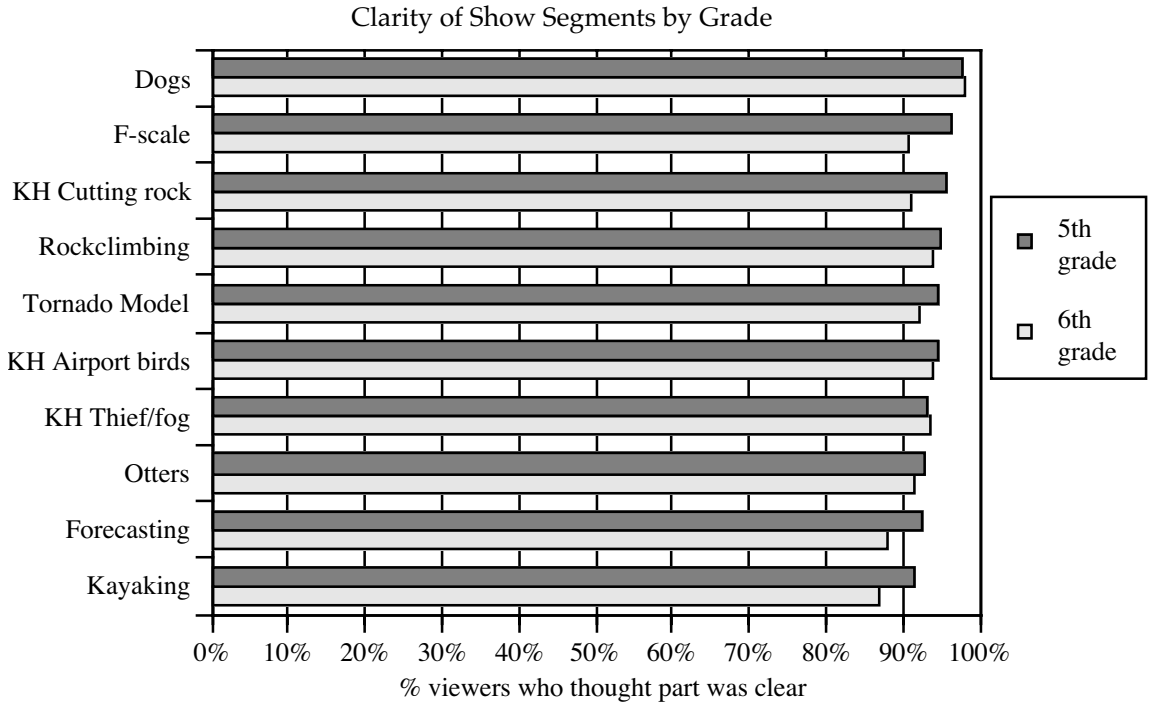
  - About 6% of the sample pointed out the importance of having kids on screen rather than adults. Sixth graders wrote the following:
    "how it had kids in every show because you can really relate to them;"
    "they had kids telling us about the topic;"
    "how they had more than just a few kids on every show;"
    "how kids were doing the experiments and not adults, it shows the process from our point of view;"
    "that kids our age are doing projects interesting to them because the subjects are confusing when adults describe it."

"when kids tried to build things, they usually made things clearer."
"way there was kids in all shows because we're kids."

- Sixth grade girls (16%) also liked the show because they learned about things in a fun way; for example:
"things they taught you how to make at home because I like making things;"
"learning new things in a fun way;"
"being able to learn science in a more exciting way;"
"most information that I did not know;"
"the interesting facts, a fun way to learn science."

Clarity ratings of show segments

Also after viewing each of the three shows, viewers rated each major segment in the show on a five-point clarity scale. The chart below presents the percent of 5th and 6th graders who were positive in their clarity ratings, saying "very clear;" "somewhat clear;" "it's ok." The segments are ordered by 5th grade percentages, which range from 98% to 91%.
The 6th grade range is 98% to 87%. More than 90% of 5th graders and 87% of 6th graders rated 10 of the 10 segments positively.



Overall, viewers felt the content of the segments was clear to them, giving clarity ratings that were as high or higher than the appeal ratings. Producers reported that project advisors were concerned that the video style and use of pop music might impair student understanding, but this appears not to be the case. Appeal and clarity ratings were significantly related for every segment, ranging from r = .43 to .68. Clarity was rated quite high by all sub-groups, but some differences still appeared between groups, as outlined in the following table.

Significant differences in mean clarity ratings for segments

Show 1 **Rocks and Minerals** (n = 295)	**Mean Clarity**	**Grade**	**Gender**	**Watch science TV or not[12]**	**Ethnicity**
Kayaking - why boulders are in rapids and not calmer waters	3.8	$5^{th} > 6^{th}$		For 5^{th} only Yes > No	
Know how - how do you cut through a rock without a saw	4.4	$5^{th} > 6^{th}$		For 5^{th} only Yes > No	For 5^{th} only Hisp > wh
Rock climbing - how do 3 types of rock affect climbing	4.0	$5^{th} > 6^{th}$		For 6^{th} only Yes > No	For 5^{th} only Hisp > wh
Show 2 **Animal Behavior** (n = 281)					
Otters - do otters groom more in the aquarium or in the wild	4.0	$5^{th} > 6^{th}$			
Know how - how do you keep birds off an airport runway	3.9			Yes > No	
Dogs - which dogs solve problems more quickly	4.6	For F only $5^{th} > 6^{th}$	F > M		
Show 3 **Weather** (n = 284)					
F-Scale - how strong was the tornado in the town of Siren	3.9				
Know how - how can weather stop a thief in their tracks	4.1			Yes > No	
Tornado model - which wind is more important - top or side	4.0	$5^{th} > 6^{th}$		Yes > No	
Forecasting weather - predict weather from nature clues	3.9		F > M		

[12] Keep in mind that differences found between those who reported watching science television and those who did not report watching are closely related to grade differences, since significantly more 5^{th} than 6^{th} graders reported watching science television (70% vs. 57%).

APPENDIX A
INTERVIEW PROTOCOL

PLEASE READ THE INTRO VERBATIM AND USE THE PROMPT QUESTIONS AS CLOSELY AS POSSIBLE TO WHAT IS SUGGESTED. DO NOT GIVE THE SUBJECTS IDEAS ABOUT HOW TO DO THE EXPERIMENT. YOU CAN ASK FOR CLARIFICATION, AS IN ASKING FOR "MORE DETAILS", BUT DO NOT HELP BY ASKING ABOUT ANY SPECIFIC METHOD OR PROCESS i.e., do not ask how many balls, how to measure, how record, etc.

Be sure to give the kids time to think. Do not move on too fast.

Other ways to work with student, without leading them on:

If student states a vague answer, parrot back the answer with a question.
S: It depends.
R: What would it depend on?

If student asks a question, parrot back the question asking how they would do it
S: How can I measure how high the balls bounce?
R: How do you think you can measure how high the balls bounce?
Only use specific process terms like "measure" if they ask a question of you. If they make a statement, for example, "I would measure how high the balls bounce", do not reply, "how would you measure."

For the pre-test, Scenario A is to be used with half your males and half your females. Use Scenario B with the other half. Then for the post-test, switch the scenarios. Those who got A for the pre-test should receive B for the post-test.

SCENARIO A

[for pre-viewing only: I'd like you to help me investigate a problem.]

[for post-viewing only: I'd like you to help me investigate a problem like the ones you saw on the *DragonflyTV* programs.]

There are no right or wrong answers to this problem. I just want to hear your ideas and what you think I could do to investigate the problem.

[read slowly] You're familiar with different kinds of bouncing balls, right? [pause] Some of them bounce better than others. I want to figure out why some balls bounce better than others. Describe what kind of experiment you might set up to find out why some balls bounce better than others.

[if S doesn't understand set up]

R: Talk to me about how you would do an experiment to find out why some balls bounce better than others.

[After S gives a first try at an idea. The following is to encourage the student to rethink their idea in more detail. You must ask it. You will need to LISTEN CLOSELY to what the student is saying and insert it into your request.]

R: I want you to imagine that you have [whatever they suggest initially - "three different sized balls" or "some balls with different amounts of rubber "]. **Now walk me through**

what you would do, step by step, to figure out why the balls bounce differently, some better than others.

[If S is vague, do not prompt with any specifics. Try some of the following prompts:]

R: Could you explain in more detail what you would do?

R: Would you explain in more detail how that would help you figure out why some balls bounce better than others?

R: Tell me more about that.

R: What would you do with the balls to find out why some bounce better than others?

[You must ask the following to get closure]

R: Can you tell me anything else that you would do to help you to come to a final conclusion why some balls bounce better?

[You must elicit review of experiment]

R: Can you think of any problems doing the experiment the way you did?
or
Any other problems that might come up?
or
Can you think of any kind of problem that might come up doing this experiment?

SCENARIO B

I'd like you to help me investigate a problem. There are no right or wrong answers to this problem. I just want to hear your ideas and what you think I could do to investigate the problem.

[read slowly] You're familiar with different kinds of jumping animals, right? [pause] Some animals jump better than others. Think about rabbits. I want to figure out why some rabbits jump better than others. Describe what kind of experiment you might set up to find out why some rabbits jump better than others.

[if S doesn't understand set up]

R: Talk to me about how you would do an experiment to find out why some rabbits jump better than others.

[After S gives a first try at an idea. The following is to encourage the student to rethink their idea in more detail. You must ask it. You will need to LISTEN CLOSELY to what the student is saying and insert it into your request.]

R: I want you to imagine that you have [whatever they suggest initially - "rabbits with different size muscles" or "big rabbits and little rabbits "]. **Now walk me through what you would do, step by step, to figure out why some rabbits jump better than others.**

[If S is vague, do not prompt with any specifics. Try some of the following prompts:]

R: Could you explain in more detail what you would do?

R: Would you explain in more detail how that would help you figure out why some rabbits jump better than others?

R: Tell me more about that.

R: What would you do with the rabbits to find out why some jump better than others?

[You must ask the following to get closure]

R: Can you tell me anything else that you would do to help you to come to a final conclusion why some rabbits jump better?

[You must elicit review of experiment]

R: Can you think of any problems doing the experiment the way you did?
or
Any other problems that might come up?
or
Can you think of any kind of problem that might come up doing this experiment?

SCORING CRITERIA FOR INTERVIEWS:

- Sets up a comparison of at least three items = 1 point
  E.g., compares performance of three different balls or three different rabbits

- Describes one or more independent variables, some characteristic on which the items vary = 1 point for one variable; 2 points for more than one variable
  E.g., balls may vary by size, composition, amount of air, weight
  E.g., rabbits might vary by size, weight, length of leg, gender

- Describes one or more dependent variables, performance measures = 1 point for one variable; 2 points for more than one variable
  E.g., bouncing performance might include height of bounce, duration of bounces, bouncing in the same place over time
  E.g., rabbit performance might be height of jump, length of jump, jumping endurance, strength of jump

- Describes some procedural steps of measuring dependent variable = 1 point
  E.g., Height of bounce - might drop balls from same height; Duration - might count number of bounces over time.
  E.g., Length of jump - might place rabbits in sand and measure distance between jumps

- Defines how to measure dependent variable = 1 point
  E.g., Height of bounce with yardstick, Duration - stopwatch and count
  E.g., Length of jump with measuring tape

- Suggests repetition of measurement/experiment = 1 point

- Describes recording of data, writing it down, charting it, dictating it, videotaping it, some method of saving the information = 1 point

- Discusses interpretation of data collected = 1 point
  E.g., the ball that has the highest inches measurement on the yardstick is the one that bounces the best and that size/composition is better
  E.g., the rabbit that jumps the farthest in the sandbox jumps the best and that size/weight/length of foot is better.

- Describes a possible problem with the experiment and suggests a solution = 1 point for problem with solution
  E.g., when asked to explain in more detail, student might observe "ball might go side-ways so I should make sure that I drop it straight down to prevent that."

A total of 11 points is possible.
Interviewers were unaware of the specific scoring criteria.
Responses were scored by Barbara Flagg.

APPENDIX B

ANALYSIS OF APPEAL & CLARITY RATINGS OF INDIVIDUAL SEGMENTS

Show 1: Rocks and Minerals

Appeal

As indicated by the overall means below, the full sample of 295 students found all three segments "somewhat" appealing. The sample liked the Know How segment most and the Kayaking and Rock Climbing segments significantly less.

How much did you like each part listed below?	Liked A lot 5	Liked Somewhat 4	It was Okay 3	Disliked Somewhat 2	Disliked A lot 1
Kayaking - why boulders are in rapids and not calmer waters		Mean = 3.5			
Know how - how do you cut through a rock without a saw		Mean = 4.0			
Rock climbing - how do 3 types of rock affect climbing		Mean = 3.5			

The **Kayaking** segment was significantly more appealing to 5th graders compared with 6th graders and significantly more appealing to kids who watched science television compared to kids who did not watch science television for both age groups. There were no appeal differences related to gender or ethnic background.

Mean appeal of Kayaking segment by subgroup

Grade		Watch Science TV	Not Watch Science TV	
5th	3.8 (n=140)	4.0 (n=97)	3.3 (n=43)	Significant
6th	3.2 (n=155)	3.4 (n=86)	3.0 (n=69)	Significant
	Significant	Significant	Significant	

The **Know How-Cutting rock** segment was more appealing to 5th graders than 6th graders as well as to those who reported watching science TV. Within 5th grade only, Hispanics liked the segment more than whites (means = 4.1 vs. 3.6) There were no appeal differences related to gender.

Mean appeal of Know How -Cutting Rock segment by subgroup

Grade		Watch Science TV	Not Watch Science TV	
5th	4.4 (n=140)	4.4 (n=97)	4.1 (n=42)	Significant
6th	3.7 (n=155)	4.0 (n=86)	3.5 (n=66)	Significant
	Significant	Significant	Significant	

The **Rock Climbing** segment was significantly more appealing to Hispanics and blacks compared with whites in the 5th grade only (means = 4.3; 4.6; 3.5, respectively). Overall, 5th graders liked the segment more than 6th graders. Within 6th grade, those who watch science TV liked the segment more. There were no appeal differences related to gender.

Mean appeal of Rock Climbing segment by subgroup

Grade		Watch Science TV	Not Watch Science TV	
5th	3.8 (n=140)	3.8 (n=97)	3.7 (n=42)	
6th	3.3 (n=155)	3.5 (n=86)	3.1 (n=66)	Significant
	Significant	Significant	Significant	

Clarity

As indicated by the overall means below, the full sample of 295 students found all three segments "somewhat clear." Correlating with appeal, the sample rated the Know How segment clearest and the Kayaking and Rock Climbing segments significantly less so.

How clear or confusing was each part listed below?	Very Clear 5	Somewhat Clear 4	It was Okay 3	Somewhat Confusing 2	Very Confusing 1
Kayaking - why boulders are in rapids and not calmer waters		Mean = 3.8			
Know how - how do you cut through a rock without a saw		Mean = 4.4			
Rock climbing - how do 3 types of rock affect climbing		Mean = 4.0			

The **Kayaking** segment was significantly clearer to 5th graders compared with 6th graders, particularly for 5th graders who say they watch science television. There were no clarity differences related to gender or ethnic background.

Mean clarity of Kayaking segment by subgroup

Grade		Watch Science TV	Not Watch Science TV	
5th	4.0 (n=140)	4.2 (n=97)	3.5 (n=42)	Significant
6th	3.8 (n=155)	3.8 (n=86)	3.6 (n=66)	
	Significant	Significant		

The **Know How-cutting rock** segment was very clear for all students, but significantly more for 5th graders who report watching science television. Also, within the 5th grade, Hispanics rated the segment clearer than white students (4.8 vs. 4.4). There were no differences for gender.

Mean appeal of KH- Cutting rock segment by subgroup

Grade		Watch Science TV	Not Watch Science TV	
5th	4.5 (n=140)	4.6 (n=97)	4.3 (n=42)	Significant
6th	4.2 (n=155)	4.3 (n=86)	4.1 (n=66)	
	Significant	Significant		

Clarity followed appeal ratings for the **Rock Climbing** segment. Hispanics rated the segment as clearer compared with whites in the 5th grade only (means = 4.5 vs. 4.1). Overall, 5th graders rated the segment as clearer than 6th graders. Within 6th grade, those who watch science TV rated the segment clearer. There were no clarity differences with respect to gender.

Mean clarity of Rock Climbing segment by subgroup

Grade		Watch Science TV	Not Watch Science TV	
5th	4.2 (n=140)	4.3 (n=97)	4.0 (n=42)	
6th	3.8 (n=155)	4.0 (n=86)	3.6 (n=66)	Significant
	Significant	Significant	Significant	

Show 2: Animal Behavior

Appeal

As indicated by the overall means below, the viewing sample of 281 students rated the Dogs segment as significantly more appealing than the Otters segment which was significantly more appealing than Know How.

How much did you like each part listed below?	Liked A lot 5	Liked Somewhat 4	It was Okay 3	Disliked Somewhat 2	Disliked A lot 1
Otters - do otters groom more in the aquarium or in the wild		Mean = 3.6			
Know how - how do you keep birds off an airport runway			Mean = 3.2		
Dogs - which dogs solve problems more quickly	Mean = 4.4				

Grade and gender played significant roles in the appeal ratings of **Otters**. Fifth graders (mean = 3.8) liked the segment more than 6th graders (3.4), girls (3.7) liked the segment significantly more than boys (3.5). Those who reported watching science television (3.7) also liked the segment significantly more than those who did not watch science TV (3.4).

Grade and gender also played significant roles in the appeal ratings of the **Know How-Airport birds** segment, as indicated in the table of means below. Fifth graders and boys rated the segment more appealing than 6th graders and girls, but 5th grade boys, in particular, liked this segment most. Those who reported watching science television also liked the segment significantly more than those who did not watch science TV (means = 3.4 vs. 3.0, respectively).

Mean appeal of Know How -Airport Birds segment by subgroup

Grade		Female	Male	
		3.1 (n=139)	3.4 (n=142)	Significant
5th	3.4 (n=132)	3.2 (n=60)	3.6 (n=72)	Significant
6th	3.1 (n=149)	2.9 (n=79)	3.2 (n=70)	
	Significant		Significant	

Although everyone rated the **Dogs** segment high in appeal, grade and gender again had a significant influence. Fifth graders (mean = 4.6) rated the segment more appealing than 6th graders (4.1). Girls (4.5) rated the segment more appealing than boys (4.2). Those who reported watching science television also liked the segment significantly more than those who did not watch science TV (means = 4.5 vs. 4.2, respectively).

Clarity

As indicated by the overall means below, the full sample of 281 students found all three segments "somewhat clear." Correlating with appeal, the sample rated the Dogs segment clearest and the Otters and Know How segments significantly less so.

How clear or confusing was each part listed below?	Very Clear 5	Somewhat Clear 4	It was Okay 3	Somewhat Confusing 2	Very Confusing 1
Otters - do otters groom more in the aquarium or in the wild		Mean = 4.0			
Know how - how do you keep birds off an airport runway		Mean = 3.9			
Dogs - which dogs solve problems more quickly	Mean = 4.6				

The **Otters** segment was rated as significantly clearer by 5th graders (mean = 4.2, n = 132) compared with 6th graders (mean = 3.9, n = 149). There were no clarity differences related to gender, ethnic background or propensity to watch science television.

The **Know How- Airport Birds** segment was significantly clearer for students who reported watching science TV (means = 4.1 vs. 3.7). There were no differences for grade, gender or ethnic background.

Dogs was very clear for all students but subgroup differences did appear, as indicated in the table of means on the next page. Girls rated the segment as clearer than boys in both grades. Fifth graders rated the segment as clearer than 6th graders, particularly for females.

Mean clarity of Dogs segment by subgroup

Grade		Female	Male	
		4.8 (n=139)	4.4 (n=142)	Significant
5th	4.7 (n=132)	4.9 (n=60)	4.6 (n=72)	Significant
6th	4.5 (n=149)	4.7 (n=79)	4.3 (n=70)	Significant
	Significant	Significant		

Show 3: Weather

Appeal

As indicated by the overall means below, the viewing sample of 284 students rated the Tornado model segment as significantly more appealing than the F-Scale and Forecasting weather segments. The Know How segment was significantly more appealing than the Forecasting segment.

How much did you like each part listed below?	Liked A lot 5	Liked Somewhat 4	It was Okay 3	Disliked Somewhat 2	Disliked A lot 1
F-Scale - how strong was the tornado in the town of Siren		Mean = 3.5			
Know how - how can weather stop a thief in their tracks		Mean = 3.6			
Tornado model - which wind is more important - the top or the side		Mean = 3.7			
Forecasting weather - can you predict weather from nature clues		Mean = 3.4			

The **F-Scale** segment was more appealing to 5th graders (mean = 3.7) than 6th graders (3.3), and more appealing to those who watch science television (means = 3.6 vs. 3.3). There were no appeal differences related to gender or ethnic background. The segment appealed more to younger viewers and those who already watch science television.

The **Know How - Thief/fog** segment was significantly higher in appeal for 5th graders, and 5th grade males compared with females (see following table). Those who reported watching science television also liked the segment significantly more than those who did not watch science TV (means = 3.8 vs. 3.2, respectively).

Mean appeal of Know How -Thief/fog segment by subgroup

Grade		Female	Male	
		3.8 (n=141)	3.3 (n=142)	Significant
5th	4.0 (n=133)	3.8 (n=61)	4.3 (n=72)	Significant
6th	3.2 (n=150)	3.0 (n=80)	3.4 (n=70)	
	Significant	Significant	Significant	

Overall, boys and 5th graders liked the **Tornado Model** segment more than girls and 6th graders; 5th grade males were the major factor in these results (see following table). Also, those who reported watching science television liked the segment significantly more than those who did not watch science TV (means = 3.9 vs. 3.5, respectively).

Mean appeal of Tornado Model segment by subgroup

Grade		Female	Male	
		3.5 (n=141)	3.9 (n=143)	Significant
5th	4.0 (n=133)	3.7 (n=61)	4.3 (n=72)	Significant
6th	3.5 (n=151)	3.4 (n=80)	3.6 (n=71)	
	Significant		Significant	

Forecasting Weather appealed significantly more to girls and 5th graders:

Mean appeal of Forecasting Weather segment by subgroup

Grade		Female	Male	
		3.7 (n=140)	3.0 (n=143)	Significant
5th	3.7 (n=131)	4.0 (n=60)	3.5 (n=71)	Significant
6th	3.1 (n=152)	3.5 (n=80)	2.6 (n=72)	Significant
	Significant	Significant	Significant	

Clarity

As indicated by the overall means below, the full sample of 284 students found all four segments "somewhat clear." The clarity mean ratings were not significantly different from each other.

How clear or confusing was each part listed below?	Very Clear 5	Somewhat Clear 4	It was Okay 3	Somewhat Confusing 2	Very Confusing 1
F-Scale - how strong was the tornado in the town of Siren		Mean = 3.9			
Know how - how can weather stop a thief in their tracks		Mean = 4.1			
Tornado model - which wind is more important - the top or the side		Mean = 4.0			
Forecasting weather - can you predict weather from nature clues		Mean = 3.9			

There were no clarity differences related to gender, grade, ethnic background or propensity to watch science television for the **F-Scale** segment.

The **Know How-Thief/fog** segment was significantly clearer for students who reported watching science TV (KH means = 4.3 vs. 3.9;). There were no differences for grade, gender or ethnic background.

The **Tornado model** segment was significantly clearer for students who reported watching science TV (4.2 vs. 3.8) and for 5^{th} (4.3) versus 6^{th} graders (3.8). There were no differences for gender or ethnic background.

Girls rated the **Forecasting** segment as clearer than boys in both grades (girls' mean = 4.2 vs. boys, 3.6).