The Physics Demonstration Show: A Force for Learning and Increasing Interest in Science?

Lenore J. Kinne and Matthew O. Zacate

Northern Kentucky University

Abstract

This study investigated the learning and affective outcomes of a polythematic physics demonstration show with topics related to the state curriculum standards and explanations appropriate for the grade level of the audience. Analysis of pre-tests administered within 24 hours before the show and post-tests administered within 24 hours after the show demonstrated modest, but statistically significant gains in both learning and interest in science. The authors recommend that providers of physics demonstration shows build on this model to measure effectiveness of their shows in terms of both learning and affective outcomes.

Keywords: physics demonstration; student learning; interest in science
The Physics Demonstration Show: A Force for Learning and Increasing Interest in Science?

Over one million\(^1\) students and adults have seen physics demonstration shows performed by physics professionals or their students in the last 40 years. The essence of a physics demonstration show is that a presenter performs for an audience a number of large, surprising, eye-catching, or otherwise remarkable demonstrations that attract attention and illustrate physics principles. Surprisingly few presenters of these shows have researched their effectiveness, either in stimulating interest in physics or in enhancing physics education. This paper describes the *Norse Physics Tour de Force*, a polythematic physics demonstration show for 4\(^{th}\)-7\(^{th}\) grade students in northern Kentucky developed with the goal of researching its effectiveness. The organization and performance of demonstration shows are considerably variable, so to help put the present work into context, we will first review characteristics of historical and current physics demonstration shows.

Physics demonstration shows vary in format. Willis and Kirwan (1976) made a distinction between a *potpourri show*, which is a collection of demonstrations chosen solely based on their wow-factor with few, if any, connections among them, and a *topic oriented show*, which is a collection of demonstrations all tied to a single topic. A search of both literature and websites yielded descriptions of 35 shows. Descriptions of four shows appeared both in a written publication and in a website (Dreiner, 2008 and Bonn University, 2008; Dahlberg, 2006 and University of Minnesota College of Science & Engineering, n.d.; Shropshire, 2009 and Idaho State University, n.d.; Sprott, 1991 and University of Wisconsin-Madison, n.d.). Of the 35 shows, we classified 21 as potpourri (Bonn University, 2008; Boone & Roth, 1992; Bryn Mawr College, n.d.; Carpineti et al., 2006; Dahlberg, 2006; Dennis, 1978; Dreiner, 2008; Hamline University, n.d; Hinko, 2010; McFarland & Kehn, 1996; Physics Factory, n.d.: Purdue University, n.d.; Rutgers University, 1999; Sarty, n.d.; Sprott, 1991; Syracuse University, n.d.; Taylor, 1996; University of California Santa Barbara, 2011; University of Minnesota College of Science & Engineering, n.d.: University of Nebraska Omaha, n.d.; University of Virginia, n.d.; Welborn, 1991; University of Wisconsin Madison, n.d.; Willey, 2010) and eleven as topic-oriented (Graham, 2003; Greenler, Lasca, Brooks & Shaw, 1993; Idaho State University, n.d.; Juwono, n.d.; Kirkpatrick & Rugheimer, 1979; Leinoff & Swan, 1993; Michigan State University, n.d.; Micklavzina, 2005; Shropshire, 2009; Shugart, 1976; University of California Irvine, n.d.; Universiteit Leiden, n.d.). Aarhus Universitet (2011) offered both topic oriented and potpourri shows. Elias (1992) also offered topic-oriented shows, but described in detail a show split into two topics, which we considered as falling between the two extremes of potpourri and topic oriented. In order to categorize such a mixed show, a new category we created a new category--a *polythematic* show, in which multiple demonstrations illustrate each of a limited number of topics (up to a maximum of four topics).

We further differentiated demonstration shows according to who presents them, audience size, and location. Physics professors, sometimes with student assistance, performed approximately two-thirds of the shows. University students, high school students, or high school

---

\(^1\)Based on tallies provided in the descriptions of shows found in the references

© 2014 Delta State University
teachers presented the remaining one-third of the shows. About 20% of shows targeted a general audience, which encompassed all ages, and the remaining 80% specifically targeted elementary, middle, or high school students. The design of twenty shows were for large audiences to view in an auditorium or gymnasium, while eight shows were for smaller, classroom size audiences and seven shows either had mixed audience sizes or did not report audience size. Most shows used the kinds of demonstrations commonly used in introductory university physics courses. Many shows had at least one large-scale demonstration targeted specifically for a large audience—the shows from the University of Wisconsin Madison (n.d.) and the University of Minnesota College of Science and Engineering (n.d.) are noteworthy for their numerous large-scale demonstrations with high thrill factor. About 30% of the shows performed at a fixed location, while about 70% of the shows traveled, usually to schools. Frequencies of shows ranged from twice a week to once a year.

Nearly all descriptions of demonstration programs emphasized that stimulating interest, generating excitement, imparting appreciation, or sharing enthusiasm in physics was the primary goal. Many concluded that their shows were effective based on audience numbers or informal feedback received following the shows (Boone & Roth, 1992; Dahlberg, 2006; Dreiner, 2008; Elias, 1992; Greenler, Lasca, Brooks & Shaw, 1993; Leinoff & Swan, 1993; Sprott, 1991; Welborn, 1991). Two programs used a survey or evaluation form completed by participants after the show to measure success (Boone & Roth, 1992; Elias, 1992), but did not specify what these instruments measured. Increases in the number, quality, and motivation of students who go on to major or minor in physics represented another indicator of success (Dennis, 1978). Few programs, however, reported explicit attempts to measure the extent to which their shows are successful. Bruce, Weissman and Novak (1997) used surveys and interviews of teachers to determine that students enjoyed the shows and benefited from exposure to real scientists. Carpineti, et al. (2006) used interviews with students in addition to feedback from teachers to conclude that their shows generated high interest.

Fewer than 20% of show descriptions indicated that communicating knowledge about physics was an important goal; nevertheless, all descriptions stated or implied that the demonstrations included explanations using underlying physics principles. To the best of our knowledge, there has been only one previous attempt to measure learning outcomes in conjunction with demonstration shows. Carpineti, et al. (2006) reported measurements of learning outcomes evaluated through teacher surveys, interviews of children, and analysis of children’s post-show drawings. Teachers reported “some improvements in the knowledge of physics of their children” (p. 909). Children who had seen the show reported better understanding of what physicists’ study, as compared with children who had not seen the show. Children’s drawings indicated memory of the demonstrations, but it appears that Carpineti, et al. (2006) did not analyze the drawings in terms of student learning gains.

A recurring theme in show descriptions is that there was a desire to keep the show fast paced and avoid detailed descriptions for fear that they would reduce interest or be unpopular. Two descriptions indicated this explicitly (Bruce, Weissman & Novak, 1997; Shugart, 1976). We inferred this to be the case for others emphasizing explanations were simple or brief (Boone & Roth, 1992; McFarland & Kehn, 1996; Sprott, 1991; Welborn, 1991; Willis & Kirwan, 1976). It appears that there is a widespread concern that inclusion of explanations that are too detailed
would cancel out the interest- and excitement-inducing benefits of the rest of the show. Although not explicitly stated in the papers cited above, it seems that a common working hypothesis is that students who have experienced interesting and exciting demonstration shows will pay closer attention in their regular science classes. Still, it would be advantageous for students to learn something from the demonstration show itself.

The work by Crouch, Fagen, Callan, and Mazur (2004) on effectiveness of demonstrations in traditional classroom instruction may provide a clue on how to structure demonstration shows so they lead to learning. Crouch, et al. investigated how demonstrations affect university-level introductory physics students’ abilities to describe and explain physical phenomena using an end-of-semester, free-response test. They found that students who actively participated either by making predictions before the performance of the show or by discussing demonstrations afterwards did better than students who only observed demonstrations and listened to accompanying explanations. This work suggests that it is important to encourage student participation in demonstration shows.

The purpose of the present study was to investigate the learning and affective outcomes of a polythematic show with topics selected from state science education standards combined with grade-level appropriate explanations. We will describe the show and the evaluation technique, based on a pre- and post-show multiple-choice test, in the method section. We will then provide results of a statistical analysis of test responses, which indicate statistically significant gains in learning outcomes and a shift toward favorable impressions toward science. Finally, we will discuss factors that likely contributed to these results and ideas for improvement.

Method

In 2005, we developed a physics demonstration show called the Norse Physics Tour de Force for performances at elementary and middle schools in northern Kentucky. Although we intended the show to be highly interesting to the student audiences, the central purpose of the show was not to entertain, but to positively affect students’ learning of basic physics concepts in particular and interest in science more generally.

Show description

Tailored to reinforce topics selected from state science curriculum standards (Kentucky Department of Education, 1999), the show included the following:

1) Materials can exist in different states and some common materials (e.g., water) can change states (Program of Studies topics S-P-PS-2 and S-4-PS-2).
2) Magnets attract and repel each other as well as certain other materials (S-P-PS-5 and S-4-PS-6).
3) Vibrating objects cause sounds (S-4-PS-5).
4) Electricity in circuits can produce light, heat, sound, and magnetic effects (S-P-PS-6 and S-4-PS-7).

---

2The name of the show is a play on words that incorporates the university’s identity, Norse, the mascot. © 2014 Delta State University
5) Matter consists of atoms, which are composed of electrons, protons, and neutrons (Core Content topic SC-H-1.1.1).

The state curriculum standards suggest introducing the above topics over a range of grade levels; nevertheless, based on interactions with students during question-and-answer segments of shows, 4th–7th grade students consistently indicated some awareness of ideas from all five topics.³

Table 1 provides an outline of the demonstration show. The three main themes of the show were electricity and magnetism, properties of matter, and sound. In addition, at the beginning of the show, we used demonstrations involving a spinning bicycle wheel (Sprott, 2006; Sutton, 1938) as a warm-up. We deemed sophisticated explanations for the bicycle-wheel demonstrations to be unsuitable for the target audience and therefore gave superficial explanations while emphasizing connections to real life. We also used the bed-of-nails demonstration to explore its benefit for explaining the concept of balanced forces, which we considered a fourth, but minor, theme of the show. We provide detailed explanations following each demonstration below.

Table 1

<table>
<thead>
<tr>
<th>Outline of the Norse Physics Tour de Force show</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theme</td>
</tr>
<tr>
<td>Warm-up</td>
</tr>
</tbody>
</table>
| Electricity and magnetism | 2, 4, and 5 | • Review of constituents of atoms and electric charge  
• Demonstrations with the Van de Graaff generator and Tesla coil  
• Review of magnetism  
• Demonstrations with the jumping ring apparatus |
| Properties of matter–part 1 | 1 | • Review of how properties of matter change with temperature  
• Demonstrations: Inflating and deflating of closed balloons |
| Balanced forces | — | • Demonstration: Lying on a bed-of-nails |
| Properties of matter – part 2 | 1 | • Demonstrations: shattering a racquetball, driving a nail into wood with a banana, jumping ring revisited |
| Sound | 3 | • Review of sound waves and introduction of resonance  
• Demonstration: breaking a beaker using sound |

³The state adopted new curriculum standards in 2006. These new standards slightly shifted the grade levels for the first introduction of topics, but all the topics selected for the demonstration shows occur in the new standards for grades 4-7.
The electricity and magnetism theme began with a question-and-answer review of the constituents of atoms and the properties of electric charge, including the characteristic that like charges repel. The presenter demonstrated the repulsion of like charges using the Van de Graaff generator to induce a volunteer’s hair to stand on end (Sprott, 2006). The presenter then introduced the ideas that electric current is a flow of negatively charged electrons and that the unit volt describes how strongly electrons are pushed. Demonstrations with a Tesla coil (Sprott, 2006) showed: (1) that an electrical arc can span the air between a metal ball on the top of the coil and a grounded metal hook and (2) that a fluorescent bulb held near the coil will light up. In the explanation of the above phenomena, the presenter asserted that the arc gave off light because it was hot, and explicitly demonstrated this by using the arc to ignite a strip of newspaper.

The electricity and magnetism theme continued with a question and answer review about magnets to remind students that magnets have poles and that likes repel and opposites attract. The presenter then introduced the idea that electric current flowing in a circle acts like a magnet in two demonstrations: (1) showing that a light bulb connected to a coil of wire placed over an electromagnet produces light, and (2) performing the jumping ring demonstration (Sprott, 2006).

The properties-of-matter theme focused on changes in state and material properties with variations in temperature. The presenter began with a question-and-answer session that focused on phase changes of water. The next demonstrations used liquid nitrogen (LN):

1) The presenter placed a small quantity of LN into a 2-liter plastic bottle, and covered the top with a deflated balloon that then inflated with nitrogen as the nitrogen underwent a liquid-to-gas transformation

2) The presenter dipped previously blown-up balloons into the LN Dewar flask to show a gas-to-liquid transformation

3) The presenter demonstrated changes in physical properties of materials by using a banana previously submerged in the LN to drive a nail into soft wood and by shattering a racquetball previously submerged in the LN (Sprott, 2006).

4) The presenter demonstrated a change in the electrical conductivity by repeating the ring jumping demonstration after submersing the ring in LN (Sprott, 2006).

It is important to ensure that the banana and racquetball remained submerged long enough to cool uniformly throughout, so the properties of matter demonstrations occurred into two parts. To ensure ample time for cooling, the presenter inserted the bed of nails demonstration at this point in the show.

The balanced forces theme came into being with the thought that the bed-of-nails demonstration could serve to help explain the concept of balanced forces. In this version of the demonstration, the presenter simply lay down on the bed of nails to show that the nails will not puncture the skin. In one variation, students saw that the nails were sharp by popping a balloon and in another variation, a co-presenter stood on a board placed across the chest of the presenter.
as he lay on the bed of nails. Instead of using the conventional explanation involving pressure (Ramsey, 2004), the presenter explained that the force of gravity acting down on the person was balanced by the forces of all the nails pushing up and with enough nails the force of each nail was small enough that they did not penetrate skin. Although impressive, we chose not to include the variation in which a co-presenter places a concrete brick on the presenter who is lying on the bed of nails and smashes it with a sledgehammer (Bucher, 1988; Hewitt, 2009) because it demonstrates a different and conceptually difficult physical principle.

The final theme on sound and vibrations potentially could lead to many demonstrations; however, there was usually time to perform just one. This theme began with a question-and-answer session about the cause of sound to remind the students that vibrating objects create sound. The presenter then asserted that sound causes objects to vibrate. Three demonstrations with a beaker in front of a loudspeaker illustrated this (Sprott, 2006).

1) A strip of paper placed on the lip of the beaker moved erratically.
2) Reflections in the beaker glass from a strobe light showed movement of the beaker’s rim in slow motion.
3) When the volume of the speaker was high enough, the beaker shattered. This demonstration served well as a grand finale of the show.

Presenters selected the equipment used in the shown in part to make it as portable as possible and in part to reduce what schools needed to provide for the show. All equipment fit on four carts, which also served as platforms to hold the demonstrations during the show. A wireless microphone and loudspeaker projected the presenter’s voice. An important component of the show was the use of a video camera in conjunction with a portable projection screen, to ensure that everyone in a large audience could see smaller demonstrations. In some cases, small demonstrations look much more impressive when magnified onto a big screen; most notably, even the arc produced by a small 50,000 V Tesla coil looks impressive on a large screen. All that was required to stage the show was a large space--preferably with controlled lighting, availability of power outlets, and stair-free access.

Method of Assessment

The show was developed and implemented in 2005, with refinements added throughout the 2005-2006 school year. Field-testing during 2005-2006 ensured that the questions used vocabulary that was both scientifically accurate and understandable to the 4th-7th grade student audience. This led to adopting identical pre- and post-tests that consisted of nine multiple choice questions on physics concepts and one multiple answer question measuring interest in science. The test questions appear in the appendix.

In 2007, we presented the show to over 2000 3rd through 8th grade students, in 11 schools. Approximately one week before each scheduled show, we sent the teachers the pre-test

---

4The main interest of this study was the show’s effectiveness in the 4th-7th grade range. One school requested that we perform the show for its 3rd grade students as well as its 4th-5th grade students. Three schools requested that their
and the identical post-test via e-mail and asked them to administer the pre-test to their students prior to the show, and hand the completed pre-tests to the presenter before the show began. We asked teachers, when possible, to administer the post-test to their students immediately after the show so they could give the completed post-tests to the presenters before they left the school building. At schools where the presentation of the show occurred at the end of the school day, we asked teachers to administer the post-tests on the subsequent day and mail them to the presenters.

In all, we collected 2,043 pre-tests and 2,008 post-tests. Of these, there was a set of 32 pre-tests from fourth graders at a school that did not contribute corresponding post-tests, and a set of 67 post-tests from seventh graders at a school that did not contribute corresponding pre-tests. We excluded these two sets of tests from the analysis, resulting in 2,011 pre-tests and 1,941 post-tests.

Results

Do students learn from physics demonstration shows?

Scoring of pre-tests and post-tests allocated one point to each correct answer with a total score of nine points possible. Table 2 provides descriptive statistics on the pre-tests and post-tests. Mean post-test scores were consistently higher than mean pre-test scores for all grade levels except the 8th grade.

To ascertain whether these learning gains were statistically significant, we conducted a 6 (grade level) x 2 (pre-test or post-test) analysis of variance (ANOVA) on the dependent variable of test score. This yielded a significant main effect of gain from pre-test to post-test, F (1, 5) = 38.28, p < .01. However, there was also a significant interaction between grade level and gain from pre-test to post-test, F (5, 3940) = 10.27, p < .01.

Table 2 follows on the next page

---

8th grade students see the show in addition to their 6th-7th grade students. Because of this, data for the 3rd and 8th grades, though limited, is included in the analysis.

© 2014 Delta State University
Table 2

Pre-test and Post-test Scores by Grade Level and School

<table>
<thead>
<tr>
<th>Grade Level</th>
<th>School</th>
<th>N</th>
<th>Pre-Test Mean</th>
<th>Pre-Test S.D.</th>
<th>Post-Test Mean</th>
<th>Post-Test S.D.</th>
<th>Gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Third Grade</td>
<td>School #9</td>
<td>77</td>
<td>3.57</td>
<td>1.49</td>
<td>75</td>
<td>4.33</td>
<td>1.63</td>
</tr>
<tr>
<td>Fourth Grade</td>
<td>School #1</td>
<td>43</td>
<td>3.79</td>
<td>1.60</td>
<td>42</td>
<td>5.29</td>
<td>1.31</td>
</tr>
<tr>
<td></td>
<td>School #4</td>
<td>121</td>
<td>3.69</td>
<td>1.64</td>
<td>116</td>
<td>5.79</td>
<td>1.45</td>
</tr>
<tr>
<td></td>
<td>School #5</td>
<td>44</td>
<td>4.84</td>
<td>1.61</td>
<td>44</td>
<td>5.43</td>
<td>1.63</td>
</tr>
<tr>
<td></td>
<td>School #6</td>
<td>67</td>
<td>3.93</td>
<td>1.62</td>
<td>65</td>
<td>5.48</td>
<td>1.70</td>
</tr>
<tr>
<td></td>
<td>School #8</td>
<td>57</td>
<td>4.14</td>
<td>1.64</td>
<td>56</td>
<td>5.75</td>
<td>1.43</td>
</tr>
<tr>
<td></td>
<td>School #9</td>
<td>60</td>
<td>4.25</td>
<td>1.55</td>
<td>43</td>
<td>5.79</td>
<td>1.28</td>
</tr>
<tr>
<td></td>
<td>School #11</td>
<td>64</td>
<td>3.91</td>
<td>1.33</td>
<td>62</td>
<td>5.92</td>
<td>1.62</td>
</tr>
<tr>
<td>All 4th grade</td>
<td></td>
<td>456</td>
<td>4.00</td>
<td>1.60</td>
<td>428</td>
<td>5.67</td>
<td>1.51</td>
</tr>
<tr>
<td>Fifth Grade</td>
<td>School #1</td>
<td>56</td>
<td>4.14</td>
<td>1.63</td>
<td>50</td>
<td>5.62</td>
<td>1.35</td>
</tr>
<tr>
<td></td>
<td>School #2</td>
<td>28</td>
<td>4.18</td>
<td>1.61</td>
<td>32</td>
<td>5.75</td>
<td>1.81</td>
</tr>
<tr>
<td></td>
<td>School #5</td>
<td>40</td>
<td>4.85</td>
<td>1.76</td>
<td>40</td>
<td>5.80</td>
<td>1.24</td>
</tr>
<tr>
<td></td>
<td>School #8</td>
<td>63</td>
<td>4.11</td>
<td>1.62</td>
<td>62</td>
<td>5.79</td>
<td>1.63</td>
</tr>
<tr>
<td></td>
<td>School #9</td>
<td>60</td>
<td>4.25</td>
<td>1.55</td>
<td>43</td>
<td>5.79</td>
<td>1.28</td>
</tr>
<tr>
<td></td>
<td>School #11</td>
<td>64</td>
<td>3.91</td>
<td>1.33</td>
<td>62</td>
<td>5.92</td>
<td>1.62</td>
</tr>
<tr>
<td>All 5th grade</td>
<td></td>
<td>328</td>
<td>4.37</td>
<td>1.59</td>
<td>325</td>
<td>5.80</td>
<td>1.55</td>
</tr>
<tr>
<td>Sixth Grade</td>
<td>School #5</td>
<td>48</td>
<td>5.42</td>
<td>1.78</td>
<td>46</td>
<td>6.35</td>
<td>1.80</td>
</tr>
<tr>
<td></td>
<td>School #7</td>
<td>88</td>
<td>4.38</td>
<td>1.82</td>
<td>114</td>
<td>5.34</td>
<td>1.87</td>
</tr>
<tr>
<td></td>
<td>School #10</td>
<td>192</td>
<td>4.80</td>
<td>1.51</td>
<td>189</td>
<td>6.22</td>
<td>1.44</td>
</tr>
<tr>
<td></td>
<td>School #11</td>
<td>57</td>
<td>4.63</td>
<td>1.83</td>
<td>58</td>
<td>6.21</td>
<td>1.55</td>
</tr>
<tr>
<td>All 6th grade</td>
<td></td>
<td>385</td>
<td>4.76</td>
<td>1.68</td>
<td>407</td>
<td>5.99</td>
<td>1.67</td>
</tr>
<tr>
<td>Seventh Grade</td>
<td>School #3</td>
<td>54</td>
<td>4.63</td>
<td>1.78</td>
<td>51</td>
<td>5.24</td>
<td>1.99</td>
</tr>
<tr>
<td></td>
<td>School #5</td>
<td>40</td>
<td>5.18</td>
<td>1.62</td>
<td>39</td>
<td>6.13</td>
<td>1.52</td>
</tr>
<tr>
<td></td>
<td>School #7</td>
<td>108</td>
<td>5.46</td>
<td>1.67</td>
<td>106</td>
<td>6.17</td>
<td>1.87</td>
</tr>
<tr>
<td></td>
<td>School #8</td>
<td>79</td>
<td>4.86</td>
<td>1.73</td>
<td>79</td>
<td>6.42</td>
<td>1.99</td>
</tr>
<tr>
<td></td>
<td>School #10</td>
<td>180</td>
<td>5.15</td>
<td>1.57</td>
<td>238</td>
<td>6.48</td>
<td>1.67</td>
</tr>
<tr>
<td>All 7th grade</td>
<td></td>
<td>461</td>
<td>5.12</td>
<td>1.66</td>
<td>513</td>
<td>6.26</td>
<td>1.82</td>
</tr>
<tr>
<td>Eighth Grade</td>
<td>School #3</td>
<td>32</td>
<td>5.13</td>
<td>2.04</td>
<td>59</td>
<td>4.68</td>
<td>2.23</td>
</tr>
<tr>
<td></td>
<td>School #7</td>
<td>107</td>
<td>5.42</td>
<td>1.49</td>
<td>108</td>
<td>6.33</td>
<td>1.87</td>
</tr>
<tr>
<td></td>
<td>School #10</td>
<td>165</td>
<td>5.18</td>
<td>1.65</td>
<td>26</td>
<td>4.92</td>
<td>2.14</td>
</tr>
<tr>
<td>All 8th grade</td>
<td></td>
<td>304</td>
<td>5.26</td>
<td>1.64</td>
<td>193</td>
<td>5.64</td>
<td>2.17</td>
</tr>
<tr>
<td>All students</td>
<td></td>
<td>2011</td>
<td>4.64</td>
<td>1.71</td>
<td>1941</td>
<td>5.86</td>
<td>1.75</td>
</tr>
</tbody>
</table>
To tease apart the significant interaction, we conducted separate bidirectional, independent samples Bonferroni t-tests at each grade level. Table 3 presents the results of these analyses. Gains from pre-test to post-test were significant at the p < .01 level for third graders through seventh graders, and at the p < .05 level for eighth graders. Estimates of effect size indicate that students in grades three through seven improved approximately one-half of a standard deviation or more from pre-test to post-test, and students in eighth grade improved approximately one-fifth of a standard deviation.

Table 3

<table>
<thead>
<tr>
<th>Level of</th>
<th>Estimated t-value</th>
<th>df</th>
<th>Significance</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>3rd grade Pre-post</td>
<td>3.01</td>
<td>148</td>
<td>&lt;.01</td>
<td>d = 0.49</td>
</tr>
<tr>
<td>4th grade Pre-post</td>
<td>15.90</td>
<td>882</td>
<td>&lt;.01</td>
<td>d = 1.07</td>
</tr>
<tr>
<td>5th grade Pre-post</td>
<td>11.60</td>
<td>651</td>
<td>&lt;.01</td>
<td>d = 0.91</td>
</tr>
<tr>
<td>6th grade Pre-post</td>
<td>10.29</td>
<td>787</td>
<td>&lt;.01</td>
<td>d = 0.68</td>
</tr>
<tr>
<td>7th grade Pre-post</td>
<td>10.23</td>
<td>972</td>
<td>&lt;.01</td>
<td>d = 0.66</td>
</tr>
<tr>
<td>8th grade Pre-post</td>
<td>2.09</td>
<td>329</td>
<td>&lt;.05</td>
<td>d = .20</td>
</tr>
</tbody>
</table>

Does students’ level of interest in science increase after watching a physics demonstration show?

The pre- and post-test measured student interest in science by one multiple response question that asked, “How do you feel about science?” It directed students to “choose as many words as you want” among the words: ‘confusing’, ‘boring’, ‘don’t know’, ‘okay’, ‘fun’, ‘interesting’. Figure 1 shows the number of times each term was chosen on the post-test as compared to the pre-test. Students consistently selected words expressing negative or ambivalent feelings toward science fewer times on the post-test, but the magnitude of the differences was small. Students consistently selected words expressing positive feelings toward science more times on the post-test, and the size of the differences was moderate. The multiple response format of the question allowed students to decide how many words to select. On the pre-test, 35 students selected all six terms, as did 50 students on the post-test. Although it is certainly possible to have mixed feelings about science, it is difficult to interpret the results of this question other than to make a broad generalization that there seemed to be a slight increase in students’ interest in science.
Learning outcomes

It was rewarding to see evidence of student learning in conjunction with the demonstration show. These results may or may not translate well to all types of shows, as the present show had three elements incorporated specifically with the aim of boosting learning outcomes. The first was selecting demonstrations that fell within only four themes rather than drawing from many areas of physics. The second was incorporating a review of topics in the form of a question and answer session within each theme to engage the audience more deeply. The third was choosing themes appropriate for the target audience by consulting the state curriculum. Still, gains in student learning, while statistically significant, were less than what one might expect given the above considerations.

Stronger gains may occur with the use of a more extensive pre- and post-show test. The use of a very short test in the present study was due to concern that a thorough test would make the whole show experience seem like a burden. Moreover, a test comprised of a small number of multiple-choice questions is relatively easy to administer and minimized the amount of time teachers needed to take away from regular class time. Because the test was short, the balance of test questions did not match the balance of the content of the show, and this may have diluted evaluation of the show’s effectiveness. Future research could address this with a split-and-switch design in which different groups of students answer different sets of test questions (Popham, 2011). The fact that students from every grade level answered the same set of
questions, which were developed to be appropriate for fourth graders, may also have reduced observed gains, particularly at the 8th grade level. It is possible that including questions with varying degrees of sophistication that align well with grade level would provide a better opportunity to distinguish learning gains in the more advanced grades.

The goal, of course, is not only to maximize observed gains but also to maximize actual gains in student learning. It seems reasonable that complex demonstrations containing sophisticated explanations would work at advanced grade levels. With a show targeted at specific grade level, a narrower range of content would enable use of a test that could be short, yet more closely balanced with the content of the show. This approach may increase learning gains. In addition, a refined pre- and post-show testing system that is more thorough than that used here would allow presenters to evaluate whether or not changes to the presentation, such as choosing different demonstrations or altering explanations, would improve students’ understanding.

**Affective outcomes**

Audiences seemed to enjoy the demonstration shows judging by the active participation of students in question-and-answer periods, applause, laughter, letters of thanks, and by direct verbal affirmation following shows. These indicators are comparable to the informal evidence used by other demonstration shows to establish success (Boone & Roth, 1992; Dahlberg, 2006; Dreiner, 2008; Elias, 1992; Greenler, Lasca, Brooks & Shaw, 1993; Leinoff & Swan, 1993; Sprott, 1991; Welborn, 1991). It is encouraging to see some quantitative evidence that supports the general perception gained from informal feedback; however, because the evidence came from responses to just one question, we must interpret these results with caution.

**Long-term versus short-term outcomes**

A notable limitation in the evaluation of learning and affective outcomes used here was that no measures of long-term effects were possible, as we asked teachers to administer the assessment immediately after the show. We adopted this method to maximize sample size over the four-month duration of this study. In initial field tests, we had asked teachers to administer post-tests and mail them to the researchers. This resulted in a low return rate, so we changed the approach to have teachers give post-tests to the presenter before he left the school.

It is unknown whether a demonstration show as an isolated event will have an enduring impact on students, no matter how interesting and exciting the show. To measure truly long-term impacts, one could use a system where teachers administer tests immediately after the show, and then again on a date well after the show. However, it would be difficult to separate the influence of the show from that of ongoing instruction.

Regardless of the effectiveness of an isolated demonstration show, good shows ideally would provide memorable events to reinforce good science teaching in the schools. As such, an attractive strategy for maximizing the long-term impact of demonstration shows is for organizers to provide follow-up material for and work with the science teachers whose students see the shows. Bruce, Weissman and Novak (1997) found that teachers, indeed, would be interested in
such an approach. Some shows are noteworthy for following such a model. Elias (1992) met
with teachers after shows to review material presented in the show, the *Physics is Fun* program
provided take-home kits with instructions so that students could repeat demonstrations on their
own (Welborn, 1991), and the UC Irvine *Physics Road Show* provides teachers with a post-show activity for one of their shows (University of California Irvine, n.d.). The Idaho State University
*Demonstration Road Show* provides the most comprehensive set of material including pre-show
teacher workshops, handouts with follow-up activities, and written guides with recommendations
how to incorporate show topics in class (Shropshire, 2009; Idaho State University, n.d.).

Learning, assessment, and interest

When designing the *Norse Physics Tour de Force* demonstration show and
accompanying assessment of its effectiveness, our primary concerns were students’ learning and
students’ interest. We decided to include detailed explanations despite concerns that such
explanations might undermine excitement generated by the demonstrations. There is likely an
optimal balance between the number of demonstrations included in a show and the amount of
time the presenter spends giving explanations. Less time spent explaining would allow for more
demonstrations, but less time spent in explanations might engender less learning.

We developed pre-show and post-show tests that were short out of a concern that
quizzing students would cause them to have unfavorable views of the show and its content.
Longer tests would require more classroom time to administer, thereby reducing classroom
teaching time, and possibly reducing students’ interest.

The present study does not provide sufficient data to draw conclusions about the optimal
balance of explanations versus numbers of demonstrations, nor about how much assessment is
too much. At a minimum, however, one can conclude that it is possible to increase knowledge,
generate favorable views, and assess the value of a show.

Implications for future research

Because of the time and the resources invested in physics demonstration shows, it is
important to establish the effectiveness of shows in terms of their impact on both learning and
affective outcomes. To establish a broader, more inclusive body of research on the efficacy of
such shows we recommend the replication of the model presented in this study. There is also
ample room for expansion of this research to explore such questions as whether some
demonstrations are more effective than others are, whether there is an optimal sequence of
demonstrations around a particular theme, whether particular explanations are more appropriate
than others, and whether any of the above considerations are dependent on grade level.

Classroom based research could explore how well gains in learning and interest are
retained over the weeks following the demonstration show and how the use of related
supplementary materials in the classroom helps sustain students’ learning and interest. If
teachers use the demonstration show as a catalyst to promote their students’ interest and follow-
up in the classroom with related experiments and deeper explanations of related topics, it is
reasonable to expect evidence of student learning gains on state test scores.
Longer-term studies also would be of benefit, as they could explore whether any increased engagement and interest resulting from shows translates into more students selecting science electives as high school courses and choosing to major in physics or other science majors in college.

Conclusion

Learning and affective outcomes of a polythematic physics demonstration show were measured using short pre- and post-show, multiple-choice tests. Results provide evidence that demonstration shows can increase knowledge and promote favorable views towards science. If evaluation in the form of tests decreases student interest, it is minimal based on the simultaneous gains in understanding and in favorable views of science observed in this study. We encourage organizers and presenters of other demonstration shows to engage in assessment of their own shows in order both to contribute to an increased understanding of how to maximize impacts of demonstration shows in general and to obtain formative data that can guide them as they seek to meet the goals of their own shows more fully.

Acknowledgements

A World Year of Physics 2005 Physics on the Road grant from the American Physical Society and grants from the Center for Integrative Natural Science and Mathematics at Northern Kentucky University provided support for this project. The authors would like to acknowledge and thank Betty Stephens for helping schedule shows, Scott Nutter, Thomas Brackman, Eric Rowley, and David McCallister for helping with transportation and presentation of shows, and the Department of Chemistry at NKU for providing liquid nitrogen.

References


Kentucky Department of Education. (1999). *Core content for science assessment V3.0 and Program of Studies*. Frankfort, KY.


Appendix

Directions: Circle the letter of the BEST answer.

1. Electricity can make:
   a. Heat
   b. Light
   c. A magnet
   d. All of the above

2. If the ends of two bar magnets attract each other, which of the following would be true?
   a. The north pole is attracting the south pole
   b. The positive charge is attracting the negative charge
   c. The two ends of the magnet have the same charge
   d. All of the above

3. If a neutral object becomes positively charged, it has
   a. More electrons
   b. Fewer electrons
   c. More neutrons
   d. Fewer neutrons

4. If you had to stay on a bed of nails for 10 minutes, which would be the best choice?
   a. Stand on one foot
   b. Stand on two feet
   c. Lie down
   d. Sit

5. If I want to carry a very heavy object, which would be the best choice?
   a. I should ask many friends to help, because the object’s weight would spread across all of the helpers
   b. I should ask one friend to help, but choose one who is very strong
   c. I should ask one friend to help, but choose one who is very trustworthy
   d. I should ask two friends to help because three people can carry most things

6. A liquid changes into a gas when
   a. Its temperature goes down a lot
   b. Its temperature goes up a lot
   c. It condenses
   d. Salt is added to the liquid

7. Which of the following is true about nitrogen?
   a. It can be a liquid
   b. It can be a gas
   c. It is part of the air we breathe
   d. All of the above

8. When sound waves hit an object
   a. The object vibrates
   b. The sound gets louder
   c. The sound gets quieter
   d. Nothing happens

© 2014 Delta State University
9. Which of the following is NOT part of an atom?
   a. Proton
   b. Neutron
   c. Molecule
   d. Electron
   e. nucleus

10. How do you feel about science? Choose as many words as you want.
    a. Confusing
    b. Boring
    c. Don’t know
    d. Okay
    e. Fun
    f. interesting