Results of the 2003 AAPT/PTRA Rural Institute Teacher Impact Study

Eric R. Banilower

December 2003

Submitted to:

George Amann
193 Primrose Hill Road
Rhinebeck, NY 12572

Warren Hein
American Association of Physics Teachers
One Physics Ellipse
College Park, MD 20740-3845

Karen Jo Matsler
3743 Hollow Creek
Arlington, TX 76001

Jan Mader
2223 5th Avenue, SW
Great Falls, MT 59404

Jim Nelson
14722 Gainesborough Court
Orlando, FL 32826

Submitted by:

Horizon Research, Inc.
326 Cloister Court
Chapel Hill, NC 27514-2296

Results of the 2003 AAPT/PTRA Teacher Impact Study

Introduction

In the summer of 2003, the AAPT/PTRA program conducted 11, week-long institutes for teachers of physics in rural schools. Seven of these institutes were sites in their first year of operation and focused their professional development on kinematics and dynamics. The professional development was intended to increase the participants' knowledge of physics content and pedagogy and to provide participants with activities they could implement in their own classrooms, with the end goal of improving student learning of physics. The logic model underlying the program's efforts is summarized in Figure 1.

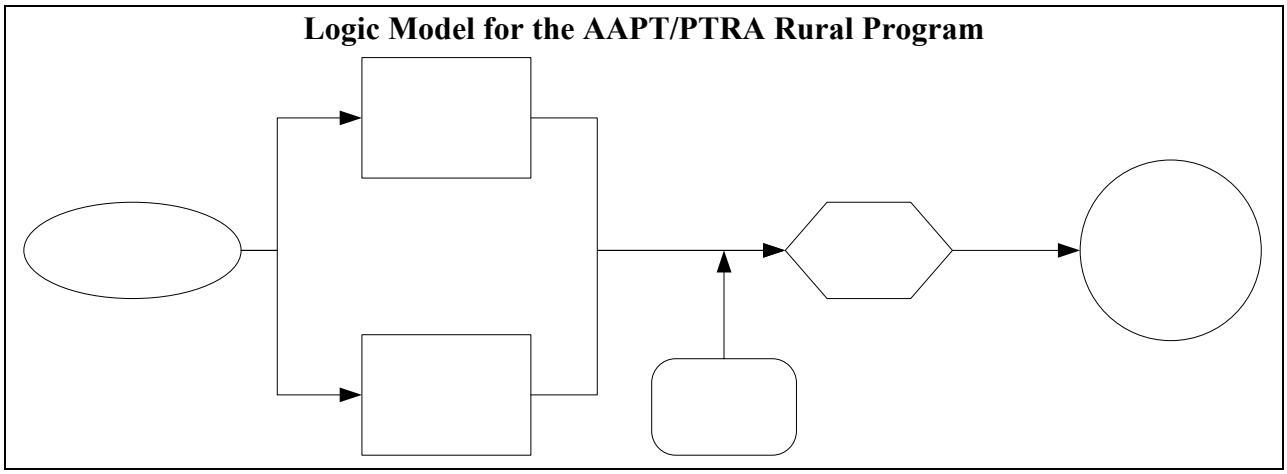

Figure 1

This study examines one of the first links in the logic model—the relationship between PTRA professional development and teacher physics content knowledge—using the results of a content assessment administered at the beginning and end of each of the seven institutes focusing on kinematics and dynamics. Specifically, this analysis seeks to answer the question, "Do teachers exhibit greater content knowledge in kinematics and dynamics after participating in a PTRA rural institute?" In addition, this study examines whether changes in teacher test scores vary by teacher gender and grade-level taught.

Instrumentation

This study employed a 35 item assessment composed of selected-response items drawn from a number of existing sources, including Jim Minstrell's *Diagnoser*. The items were selected, with the assistance of the project leadership, based upon the content goals of the rural institutes. The assessment targeted common concepts in kinematics (the description of motion) and dynamics (the effect of forces on motion). A copy of the assessment can be found in Appendix A.

The assessment yields two scale scores, one for kinematics and one for dynamics, each of which is composed of 17 items.[1] Each scale score is computed as the percent of items correct. Table 1 shows the number of items and reliability (Cronbach's alpha) for the assessment scales; each scale has an acceptable reliability[2] indicating the items within each set are well correlated with each other and appear to be measuring the same construct (e.g., kinematics knowledge). In addition, teacher background data (e.g., teacher demographics) from a questionnaire completed by the participants at the beginning of the rural institutes were also used in this study.

Table 1
Assessment Scale Reliabilities

	Number of Items	Reliability (Cronbach's Alpha)	
		Pre	Post
Overall	34	0.86	0.82
Kinematics	17	0.76	0.72
Dynamics	17	0.82	0.74

The leaders of each rural institute were asked to indicate for each item on the assessment whether or not they would expect participants to be able to answer the item correctly after participating in their institute. Table 2 shows their responses. Of the 34 items used in these analyses, 30 items received affirmative responses from at least 4 of the 7 institutes; only four items received three or fewer affirmatives.[3] However, these four items were retained in the analysis since the project leadership indicated they covered content targeted by the project. Overall, these data indicate that the assessment was fairly well aligned with the content goals of the rural institutes.

[1] Q1-Q17 focused on kinematics, Q18-Q35 focused on dynamics. One item, Q34, was dropped from all analyses due to a flaw in the item.

[2] Typically, a Cronbach's alpha ≥ 0.60 is considered acceptable, ≥ 0.70 is fair, ≥ 0.80 is good, and ≥ 0.90 is excellent.

[3] The analysis was conducted without Q28 as well, but yielded similar results.

Table 2
Institutes Where Leader Indicated Assessment Item Should Be Sensitive to their Professional Development

Item	Number of Institutes (N=7)
Q1	7
Q2	6
Q3	7
Q4	6
Q5	7
Q6	7
Q7	7
Q8	6
Q9	7
Q10	6
Q11	6
Q12	5
Q13	7
Q14	4
Q15	3
Q16	6
Q17	4
Q18	3
Q19	7
Q20	3
Q21	7
Q22	6
Q23	4
Q24	6
Q25	7
Q26	5
Q27	6
Q28	1
Q29	5
Q30	6
Q31	6
Q32	6
Q33	6
Q35	6

The Sample

The assessment was administered at the beginning and end of each of the seven rural institutes that focused on kinematics and dynamics. HRI received complete responses from 150 out of 159 rural institute participants, a response rate of 94 percent. Table 3 shows the demographic characteristics of these participants. Just over half of the participants were male; nearly all were white. The participants have a wide range of teaching experience, with roughly half having taught for more than 10 years. Eighty-five percent of the participants taught at the high school level, 15 percent taught in the middle or elementary grades.

Table 3
Demographics of Participants with Complete Data

	Percent of Respondents
Gender	
Male	55
Female	45
Race/Ethnicity	
American Indian/Alaskan Native	1
Asian	1
Black or African American	0
Hispanic or Latino	1
Native Hawaiian or Other Pacific Islander	0
White	96
No Response	1
Prior Teaching Experience	
0–2 Years	11
3–5 Years	14
6–10 Years	21
11–15 Years	20
16–20 Years	17
21–25 Years	11
26 or more Years	7
Grade Level Taught	
Elementary School	1
Middle School	14
High School	85

Analysis and Results

Descriptive statistics for the pre- and post-test scores are shown in Table 4. Overall, teachers appear to have scored higher on the post-test than on the pre-test. Teachers also appear to have scored higher on the kinematics sub-scale compared to the dynamics sub-scale. A repeated measures analysis of variance model was used to statistically test changes in teachers' assessment scores. Individual item statistics are presented in Appendix B.

Table 4
Descriptive Statistics for the Teacher Assessment

	Minimum	Maximum	Mean	Standard Deviation
Pre-Test				
Overall	26.47	97.06	69.08	17.58
Kinematics	23.53	100.00	73.57	18.80
Dynamics	5.88	100.00	64.59	21.09
Post-Test				
Overall	32.35	100.00	73.84	15.12
Kinematics	23.53	100.00	77.92	16.75
Dynamics	29.41	100.00	69.76	17.63

Teacher gender and teaching assignment[4] were also included in the analyses to see if performance was consistent across different types of participants. Table 5 provides a summary of the results. These results are described more fully in the following sections.

Table 5
Results by Research Question†

	Overall	Kinematics	Dynamics
1. Did participants score higher on the post-test than on the pre-test?	Yes	Yes	Yes
2. Did changes in test scores vary by gender?	No	No	No
3. Did changes in test scores vary by grade-level?	Yes	Yes	No
4. Did changes in test scores vary across gender/grade-level combinations?	No	No	Yes

† All "Yes" results are statistically significant at the $p < 0.05$ level.

Research Question 1: Did participants score higher on the post-test than on the pre-test?
Participants in the AAPT/PTRA rural institutes scored significantly higher on the post-test than on the pre-test. On average, post-test scores were 5 percentage points higher than pre-test scores. Further, these results are consistent across the two sub-scales, indicating the project had an impact in both content areas. However, mean post-test scores around 70 indicate that the participants had not mastered all of the content goals set forth by the project by the end of the rural institutes. One way to think about the size of the impact is to consider the average number of items participants answered correctly on the pre- and post-tests. On average, participants answered about one additional item correctly on each sub-scale, or two items overall, on the post-test than they did on the pre-test. (See Table 6.)

Table 6
Mean Number of Items Answered Correctly

Scale	Number of Items	Pre	Post
Overall	34	23.49	25.11
Kinematics	17	12.51	13.25
Dynamics	17	10.98	11.86

[4] Because of the small number of elementary teachers in the sample, teachers were categorized as either "high school" or "elementary/middle school."

It's important to note, however, that the project planned on holding two day-long follow-up sessions during the school year to continue working with teachers. The post-test was administered on the last day of the summer institute in order to maximize response rates. Thus, these results should be interpreted as the impact of the AAPT/PTRA rural summer institute, not the entire AAPT/PTRA program. It is possible that post-test scores would have been higher if the assessment had been administered after the two follow-up sessions.

Research Question 2: Did changes in test scores vary by gender?

Figure 2 shows the mean pre- and post-test scores for male and female participants. Female participants scored lower on the pre-test than did male participants (mean scores of 61 for females and 68 for males), due mainly to the scores on the dynamics sub-scale (mean scores of 55 for females and 64 for males). Despite the initial disparities, male and female participants exhibited statistically equal amounts of growth, both overall and on the two sub-scales. As a result, female participants scored significantly lower on the post-test than did male participants (mean scores of 67 for females and 78 for males), again due mainly to performance on the dynamics sub-scale (mean scores of 60 for females and 75 for males).

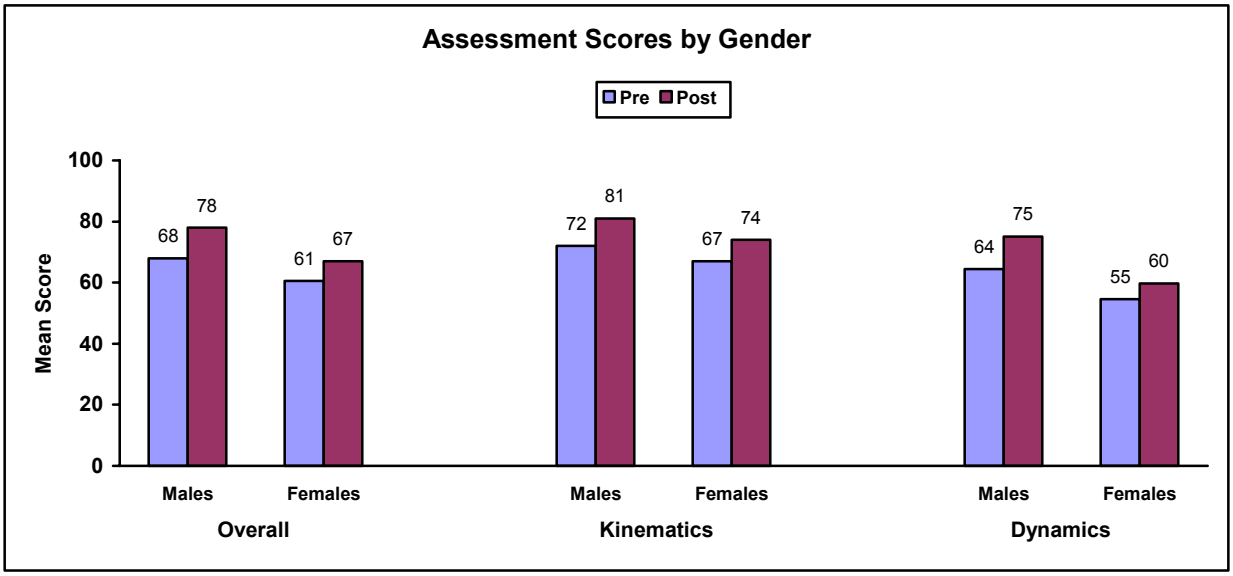

Figure 2

Research Question 3: Did changes in test scores vary by grade-level?

As can be seen in Figure 3, high school teachers entered the rural institutes with stronger content backgrounds than elementary and middle school teachers (average pre-test scores of 72 and 58 percent, respectively). However, elementary and middle school teachers exhibited statistically larger gains on the assessment, effectively narrowing the gap on the kinematics sub-scale and the overall test score with average gains of 12 points for elementary/middle school teachers and 3 points for high school teachers on the overall assessment.

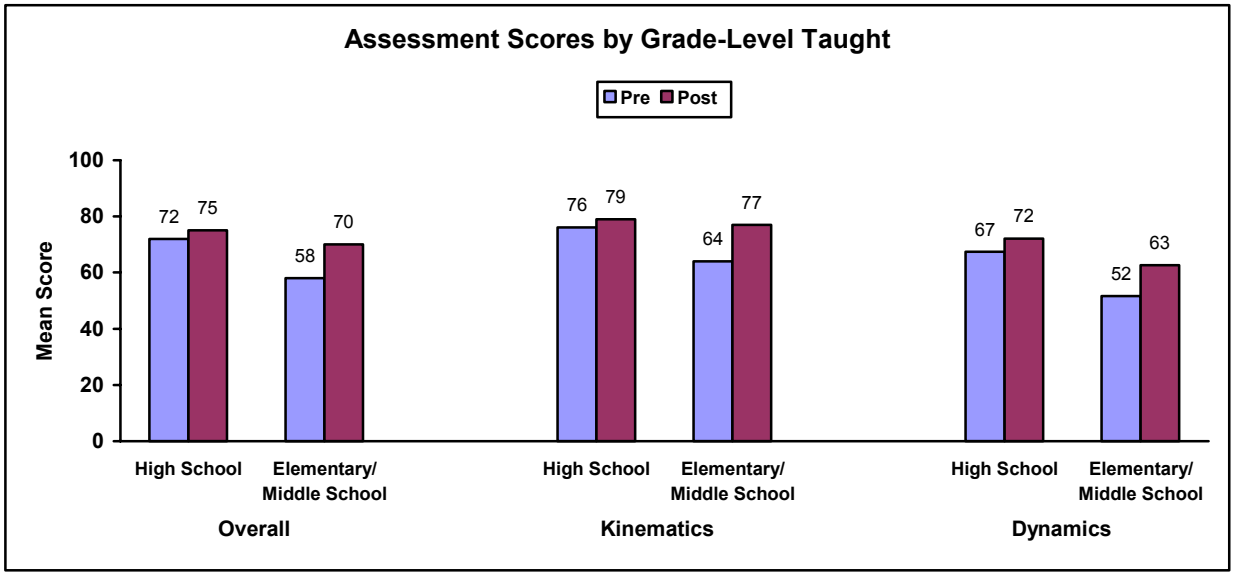

Figure 3

Although elementary and middle school teachers appear to have made more progress in dynamics than did high school teachers, this result is not statistically significant, due in part to the larger standard deviation for the dynamics sub-scale.

Research Question 4: Did changes in test scores vary across gender/grade-level combinations?

The data were further disaggregated by gender and grade-level combinations, creating four sub-groups: high school males, high school females, elementary/middle school males, and elementary/middle school females. (See Figure 4.) On the kinematics sub-scale and the overall assessment score, no statistically significant difference was found in the performance of these sub-groups, indicating that the amount of growth for each group was similar. However, a significant difference in group performance was found on the dynamics sub-scale.

Examining the results on the dynamics sub-scale, two findings are apparent. One is that male elementary/middle school teachers had much larger gains on the dynamics sub-scale than any other group (18 points compared to an average of 5 points for the other groups). Second, female elementary/middle school teachers had much lower scores on the pre- and the post-test than any other group.

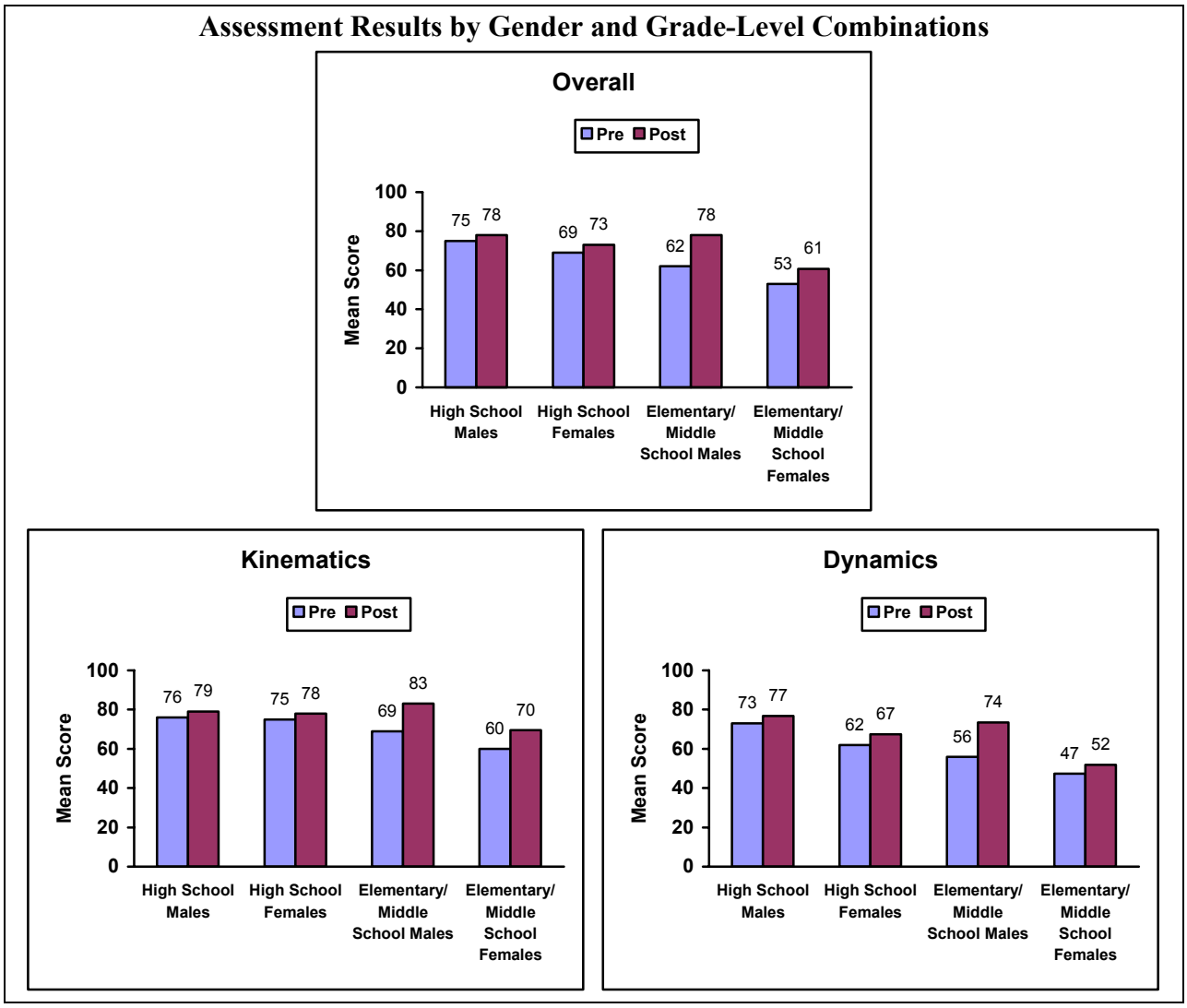

Figure 4

Conclusions

The results of this study provide preliminary evidence of the impact of the AAPT/PTRA rural program on teacher's physics content knowledge. The rural institutes appear to have had a positive impact on teachers' knowledge of kinematics and dynamics. These impacts are found for both male and female teachers. In addition, gains made by elementary/middle school teachers are larger than those made by high school teachers.

Additionally, these data provide the project an opportunity to reflect on its efforts and to make any needed adjustments for future rural institutes. Some questions the project may want to consider are:

1. Is a mean post-test score of 74 percent on this particular assessment good enough? If not, what can be done to increase the learning of all teachers? Although the project had goals

other than content related ones for the rural institutes, is the project satisfied with a 5 percentage point increase (equivalent to two items on the overall assessment) in scores?

2. Are elementary/middle school teachers making larger content knowledge gains than high school teachers because they have more room to grow, or are the institutes teaching to the lowest common denominator? What can be done to increase the likelihood of growth for all teachers?

3. Given the historical inequities in the physical sciences, and the fact that 90 percent of elementary school teachers, 70 percent of middle school teachers, and 48 percent of high school teachers in rural areas are female,[5] are there additional steps the project should take to close the gap in scores between male and female teachers? In addition, the project may want to examine why they appear to be more successful working with male elementary/middle school teachers than female elementary/middle school teachers.

[5] All comparisons to the national science teaching force are based on special tabulations of data from *The 2000 National Survey of Science and Mathematics Education*.

Appendix A

RPTRA Kinematics and Dynamics
Teacher Assessment

Rural PTRA
Kinematics and Dynamics

Dear Colleague: Thank you for your cooperation in completing this task. The PTRA project is requesting this information to assess its progress in achieving the program's goals, and to provide data to the National Science Foundation, the funding agency for the project. Please be assured that your responses will be treated with the strictest confidentiality in that only group results, utilizing data from this and six other rural institutes, will be reported. Again, thank you for your assistance.

Instructions: Mark your answers on the answer sheet. Fill in only one circle for each question. Please do not write on this test booklet; you may use scrap paper if needed. Please do your best, but don't be concerned if you do not know the answers to all of the questions at this time. The project hopes that your knowledge of these concepts will increase over the course of your participation.

Items adapted with permission from Jim Minstrell's *Diagnoser*.

For information contact:

Horizon Research, Inc.
326 Cloister Court
Chapel Hill, NC 27514
Email: hri@horizon-research.com

Appendix B

Item Statistics