ACTION RESEARCH

EVALUATION OF *MATERIALS WORLD MODULES:* 2006 SUMMER INSTITUTE

conducted at Garrett College
in Deep Creek Lake, Maryland

prepared for

Building Engineering and Science Talent and the U.S. Department of Defense

BEST
Building Engineering
& Science Talent

ACTION RESEARCH & ASSOCIATES, INC.
P.O. Box 494, Ellicott City, MD 21041-0494
drkjuffer@comcast.net   www.actionresearchinc.com

(This page intentionally blank)

EVALUATION OF *MATERIALS WORLD MODULES*
PHASE 1: 2006 SUMMER INSTITUTE

TABLE OF CONTENTS

Table of Contents (continued)

Tables and Figures

(This page intentionally blank)

EXECUTIVE SUMMARY


Students creating polymers


Students designing a new concrete product


Students testing composites

Evaluation Objectives

As the first phase of a multi-year plan to measure the effectiveness of *Materials World Modules* (*MWM*), the U.S. Department of Defense (DoD) sponsored a Summer Institute for 82 middle and high school students at Garrett College, in Deep Creek Lake, Maryland, from July 9 - August 5, 2006. The Summer Institute was a residential immersion science and math institute designed to explore the relative effectiveness of a traditional teacher-directed science curriculum and textbook compared to the *Materials World Module*s materials and methods, an inquiry- and design-based materials science program.

DoD has made a multi-year commitment to support the scale-up of *MWM* in areas that host a significant concentration of DoD research facilities. A rigorous evaluation of *MWM* was deemed essential to the success of its dissemination and the development of a national rollout model. DoD selected Garrett College to implement this plan.

Employing a multi-method, quasi-experimental design with random assignment of students to Treatment and Comparison Groups in matched pairs, the evaluation of the Summer Institute had three main purposes:

- To gather preliminary evaluation data to inform decisions of stakeholders about *MWM*;
- To develop reliable assessment instruments that can be used in a more comprehensive classroom evaluation in Maryland and elsewhere; and
- To provide formative feedback to guide the training of teachers to teach *MWM*.

Description of the Intervention

During the Summer Institute every effort was taken to control key aspects of the students' environment during the four-week, 24-hour, seven-day Summer Institute, every effort was taken to ensure that only the independent variable of interest – *MWM* versus traditional classroom science methods and text – varied between Treatment and Comparison Groups. The Summer Institute focused on four science units - Composites, Sports Materials, Concrete and Polymers. One unit was taught each week for three hours per day Monday through Friday.

Prior to the Summer Institute, a curriculum committee developed common "Learning Objectives" for each unit, delineating key science principles to be covered in each unit by both groups. The Treatment Group was taught with a predominantly student-centered, hands-on approach using *MWM* reading materials and inquiry-based hands-on laboratory experiences combining data collection, design challenges, and interactive class presentations. The Comparison Group was taught science in a traditional classroom style, with predominantly teacher-centered lectures, teacher-directed lab experiences, and a science textbook (Pearson Prentice Hall's *Physical Science: Concepts in Action: With Earth and Space Science*).

A weekly incentive program replaced the motivating factor of earning "grades" for academic work. The combination of 1) physical separation of the treatment groups during the day; 2) classroom groups assigned neutral names like "Platinum" and "Silver;" and 3) quality teaching of the same content areas, irrespective of group assignment, resulted in students not discovering which treatment group they had been assigned to until the final week of the study.

Sample Development and Selection

The Summer Institute recruited from a pool of more than 400 middle and high school students in order to fill 96 spots. To develop this pool, in April 2006, Garrett College sent professionally-designed program announcements and posters to the Maryland Science Teachers Association and to every Maryland school district via superintendents, principals, science curriculum coordinators, and department chairs. The materials invited the students to apply for the Institute via the Internet. Financial cost was not a barrier to student participation because DoD paid the expenses for all selected students, enabling students from a wide range of socio-economic backgrounds to attend. Students came from virtually every county across Maryland, from inner-city schools, private schools, parochial schools, charter schools, large public schools in suburban and urban areas, military academies, technical schools, and "home schools." As a result, it was possible to recruit and select a diverse cross-section of students from every part of the state.

A two-tiered random sampling method was developed to first select the participants, and then assign them to a treatment group in order to create a probability sample. The sampling model reflected Maryland's demographics regarding race and sex, adjusted proportionately to a sample of 96 participants. In the final sample, of the 82 students who completed the four-week Institute, the demographic factors were evenly distributed between the Treatment and Comparison Groups. Sample attrition (due to first day "no shows," home sickness, discipline issues, and family vacations) was distributed evenly between the two treatment groups. Both groups had 13 high school and 28 middle school students. The proportion of male and female students was evenly split, with 41 of each, and proportionately split across middle and high school levels. Proportions of Black, White, and "Other" students in the final sample were 33%, 60%, and 7%, respectively, comparable to the U.S. Census Bureau's statistics for Maryland.

Evaluation Instruments

Action Research & Associates, Inc., an independent research firm with test development expertise, developed a series of data collection instruments to measure cognitive and attitudinal changes in students. The data from the students were triangulated against the perceptions of their teachers and with the independent researcher observations. The scientifically-developed instruments, supplemented by classroom observations, included:

Pre- and Post-Tests: Weekly tests administered before and after each module which measured gains in student content knowledge. Tests consisted of 20-33 items each, with about half from *MWM* and about half from the Pearson Prentice Hall text. The resulting combined tests had a respectably high estimated reliability of .93 (coefficient alpha).

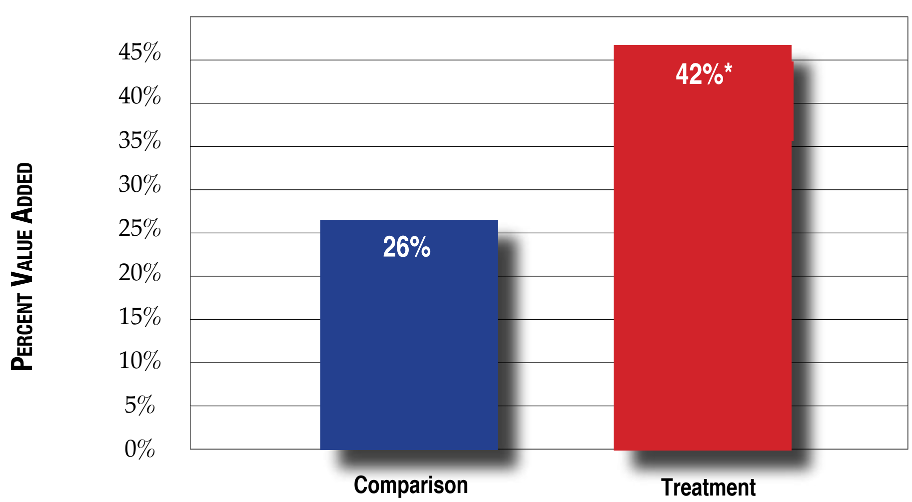
Pre- and Post-Institute Student Surveys: The survey instruments captured the students' self-assessed changes in the following: 1) their attitudes toward science and careers; 2) their science skills; 3) their team-work skills; and 4) their ability to use scientific inquiry to problem solve in the classroom before and after the Institute.

Teacher Post-Module Surveys: The Control and Treatment teachers provided weekly information on their perceptions of changes in the students' attitudes, science and inquiry skills and ability to problem solve in the classroom.

Results

The evaluation of the Summer Institute measured the cognitive and attitudinal gains of a randomly selected, small, yet diverse sample of students in a controlled environment. The objectives were 1) to assess the potential benefits of *MWM* relative to traditional classroom instruction, and 2) to prepare for a full-scale evaluation that will measure additional variables. With this qualification in mind, the following preliminary findings stand out:

Figure 1
Gains in Science Knowledge by Treatment Group



Base: Pre-Test Scores

* statistically significant at $p<0.0001$

Students' Science Knowledge Gains

Control versus Treatment

Summer science immersion programs produce noteworthy gains for all students, but ***MWM* students made significantly larger knowledge gains than those in the Comparison Group for the same time investment.** (Figure 1, previous page).

Due to the two-tiered randomized selection process, Control and Treatment students achieved the same average scores on the baseline pre-tests. By the end of the Institute, Treatment and Comparison students both made appreciable gains in science knowledge. However, the Treatment Group improved their pre- to post-test science knowledge scores by an average 42% Percent Value Added relative to the Comparison Group, which averaged a 26% gain.

This result represents a statistically significant ($p<0.0001$) difference, suggesting a strong probability that the treatment (*MWM*) was more effective than traditional classroom approaches in helping students learn more science.

Demographic Groups

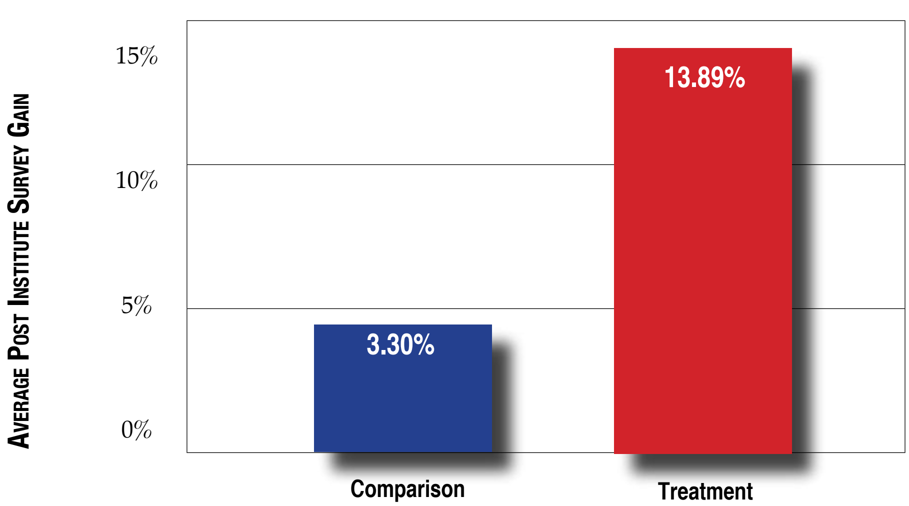
The gains of the Treatment Group students varied across demographic groups.

Both girls and boys taught with *MWM* out performed their Comparison Group peers, with **Treatment girls gaining an average of twice as much science knowledge as Control Group girls**.

White students achieved a 44% Percent Value Added compared to their Comparison Group counterparts' 27% gains in science knowledge. **Non-White Treatment students (37%) out-gained the Comparison Group Non-White students (25%) as well as the Comparison White students (27%)**.

Both middle and high school Treatment students (41% respectively) out performed the Comparison Group middle and high school students (29%, 31% respectively).

FIGURE 2
STUDENTS' SELF-ASSESSED GAINS IN ATTITUDES AND SCIENCE SKILLS



BASE: PRE-INSTITUTE SURVEY SCORES

Student-Reported Attitudinal Changes

Treatment and Comparison students entered the Summer Institute with similar self-assessed attitudes toward science as well as similar science skills. Due to the two-tiered randomized selection process, a pre-Institute battery of more than 50 items produced virtually the same average results for Treatment and Comparison students.

The items measured students' self-assessed favorability toward science study and career aspirations, use of inquiry and problem-solving, and teamwork skills.

After receiving the same amount of instruction, Treatment students learning with *MWM* reported more improvement in their attitudes toward science and in science skills than did Comparison students over the course of the Institute.

The two Groups' average positive change (13.89 - Treatment vs. 3.30 - Comparison) approaches a level of statistical significance ($p<0.07$) (Figure 2, previous page).

Table 1: Students' Perceived Gains in Attitudes and Science Skills-Selected Items

Attitudes, Skills, Behaviors	*Comparison*	*Treatment*	*t-test*
Attitudes towards Science Scale	**0.25**	**2.45**	**0.12**
Science classes are interesting.	-0.28	0.24	0.02
Laboratory science is boring.	0.30	-0.18	0.04
I enjoy doing science experiments.	-0.08	0.34	0.04
Inquiry Skills and Problem Solving Scale	**0.80**	**3.68**	**0.05**
My teacher asks questions to stimulate me to come up with my own answers.	-0.13	0.50	0.01
Science classes encourage me to discuss my ideas.	-0.08	0.45	0.05
I develop scientific explanations following rules of logic and evidence.	-0.20	0.29	0.03
I provide alternate explanations to solve a problem.	-0.18	0.21	0.06
I apply concepts/ideas I've learned in real-world design problems	0.08	0.63	0.03
I design useful things in science class.	0.13	0.55	0.04
Design a test of the product/project or lab	0.15	0.61	0.07
Team Skills and Personal Improvement Scale	**-0.65**	**0.74**	**0.07**
Be more inquisitive	-0.35	0.08	0.04
Be more self-reliant and take charge of my own learning	-0.05	0.32	0.06
Total Scale *(53 items)*	**3.30**	**13.89**	**0.07**

Conclusion

Since the evaluation results from the Summer Institute are based on a small test sample, the findings will be evaluated and validated with larger student samples taught in standard school-year educational environments between 2007 and 2009 in Maryland. Nonetheless, the findings from the Summer Institute evaluation provide a preliminary indicator of the *Materials World Modules'* potential to positively impact student science learning via inquiry- and design-based instruction. Commencing in 2007, the Department of Defense is funding a pilot program to support the use of inquiry- and design-based instruction through *MWM* and other instructional materials in the State of Maryland and nationwide.

(This page intentionally blank)

ACTION RESEARCH

Evaluation of *Materials World Modules:* 2006 Summer Institute

conducted at Garrett College
in Deep Creek Lake, Maryland

prepared for

Building Engineering and Science Talent and the U.S. Department of Defense

BEST Building Engineering & Science Talent

Action Research & Associates, Inc.
P.O. Box 494, Ellicott City, MD 21041-0494
drkjuffer@comcast.net www.actionresearchinc.com

Background of Study

Overview of *Materials World Modules*

MWM's Development

Materials World Modules (MWM) (www.materialsworldmodules.org) is a hands-on, interdisciplinary approach to teaching science that provides students opportunities to learn science and math in the context of real-world applications. Because the world of materials covers all materials, both natural and synthetic, in every aspect of our lives, the study of materials makes it possible to integrate multiple disciplines while also connecting textbook theory, hands-on application, and inquiry-based design.

Each *MWM* module brings together tools in science, technology, and mathematics in a 10-hour block of teacher-guided, student-centered instruction. The modules have flexible formats allowing them to supplement existing curricula in earth science, biology, chemistry, physical science and mathematics classes. They can be adapted to STEM curricula at the middle and high school levels and to collegiate studies at the introductory level.

MWM was developed by a research team led by Dr. Robert P.H. Chang, Professor of Materials Science & Engineering and Professor of Electrical & Computer Engineering at Northwestern University. In 1994, Dr. Chang received a three-year $1.8 million dollar grant from the National Science Foundation to develop educational science modules for middle and high schools to improve U.S. science education through interdisciplinary teaching and learning. In 2001, the program received another $3.6 million from the National Science Foundation and the State of Illinois.

The *MWM* Program is the product of more than 14 years of research, development and refinement. During 2002, students in 37 states participated in field-testing the program. Additional program development and research with students was conducted in 2003. The program has initiated web-based dissemination of its curriculum materials and teacher-training program with the goal of evolving *MWM* into an Internet-based global science education program.

MWM's Rationale

The developers of *MWM* asked science teachers what they needed to 1) improve students' inquiry skills and problem-solving capabilities, 2) increase their scientific and technological literacy, 3) provide real-world applications of math and science, and 4) promote integration of subject matter across disciplines.

Drawing heavily upon teacher input, the developers produced a series of inquiry- and design-based science modules designed to be dropped into an existing science curriculum, each lasting approximately 10-12 instructional hours. These modules were meant to serve as curriculum enhancements that could be flexibly used to enrich curricula in the physical and life sciences.

MWM uses materials science and hands-on labs and design projects to interest, motivate and engage students in science. Each module is centered around a *design project*, which encourages students to learn science and math through their application to real-world problems. The activities in each module are designed to provide the necessary background and motivation for students to successfully complete the culminating design project. Each module activity builds on preceding ones and ultimately provides students with the principles, ideas, and techniques that they will use in the design challenge.

Learning by Inquiry through Design: The modules incorporate the notion of inquiry through design. These two ideas are thought to be important to the learning of science. They involve participation in the process of scientific inquiry and hands-on application of scientific principles through engineering design.

The theory behind *Materials World Modules* is that students need to understand that science is an ongoing search for better and better explanations, rather than a collection of sanctioned facts. To this end, students need to participate in the process of finding explanations for phenomena they find interesting. Engaging in scientific theory involves finding a question for which one wants to know the answer. The process used to answer the questions usually involves experimentation or research. In *MWM*, students design the investigation. The *MWM* text materials and activities are designed to help the students organize their ideas and understand the relationship between their original question and the results of their investigations.

In science, student scientists not only must learn how to engage in scientific inquiry, but also must be able to evaluate the answers that they produce. That is, scientists must be able to convince their peers that the results of their investigations provide the best explanation. This is an often overlooked aspect of scientific investigation in the classroom. In *MWM*, design projects culminate in class presentations where students argue for the merits of their own designs, while critically examining others' designs.

Learning by Iterative Design: *MWM* incorporates a design approach called *iterative design* consisting of a series of cycles that involve designing a product, testing it, and then redesigning it based on what is learned from the previous design. Each module is centered around a design project, which encourages students to learn science and math through their application to real world problems. The activities in each module are designed to provide the necessary background and motivation for students to successfully complete the design project presented. Each module activity builds on preceding ones and ultimately provides students with the principles, ideas, and techniques that they will use in the design challenge.

This approach is formulated from the real-world example that few engineers or scientists get something right the first time. They attempt to create improved designs each time they iterate through the design cycle. In the classroom, *iterative design* provides the students the opportunity to learn something about their design, and then apply that knowledge to improve it to produce results.

This approach is designed to produce two main benefits. First, students will be highly motivated because their design ideas will be drawn from real-world experience, not just from science class. Second, the design process will allow students to appreciate how science works in their daily lives.

MWM's Learning Goals

The development of the *MWM* modules is based on Learning Goals that are consistent with the National Science Education Standards (NSES). The Learning Goals include the following:

- **Develop the abilities necessary to do scientific inquiry**: Including the ability to generate questions, design and conduct scientific investigations, formulate models, analyze alternative models, and communicate and defend explanations.
- **Understand scientific inquiry:** Understanding that scientific inquiry is focused on logically consistent explanations, grounded in current knowledge, and augmented by mathematics and technology.
- **Become familiar with materials science:** Developing an understanding of materials science by applying knowledge from the physical sciences, as well as life and earth sciences to create materials for specific purposes.
- **Take part in iterative design:** Providing opportunities to address technological problems, propose designs, choose between alternative solutions, implement and evaluate a solution, redesign the product, and describe in written and oral forms the experimental design, process, and results.
- **Understand the relationship between science and technology:** Understanding the differences between the purposes and nature of scientific and technological studies and the interrelationships between these fields.
- **Understand contemporary problems**: Appreciating the use of science and technology to meet local, national, and global challenges including problems of personal and community health, natural resources, environmental quality and human-induced hazards.
- **Acquire a historical perspective:** By examining the history and the nature of science as a human endeavor, produces new knowledge supported by developing technology.

Capsule Descriptions of MWM Modules

The 2006 Summer Institute taught the Composites, Concrete, Sports Materials, and Polymers Modules (marked with italics below). The nine *MWM* modules currently available include the following:

Biodegradable Materials Module: Students make, test and evaluate biodegradable films and gels. They use their knowledge to design devices that release a dye in a controlled manner as the materials degrade.

Biosensors Module: Students investigate the use of biological molecules as materials and use enzymes as chemical sensors in the design of diagnostic tests for peroxide, glucose and cholesterol.

Ceramics Module: Students study the science of compacting ultra-small ceramic particles. They evaluate the evolution of density and microstructure of ceramics as they are synthesized at high temperatures. They then use ceramics to make a voltage-protecting device.

Composites Module* Students find out what composite materials are and test them to learn the advantages over pure materials. They design a prototype composite material to make a strong, lightweight fishing pole.

Concrete: An Infrastructure Material Module: Students learn how the components of concrete can be modified to alter its properties. They use their knowledge to make concrete roofing tiles that meet specific design and performance criteria.

Food Packaging Module: Students learn about the many functions of food packaging, including protecting foods and packaging materials without harming the environment. Then they design their own environmentally friendly package for delivering a hot baked potato.

Polymers Module: Students examine the viscoelastic, mechanical and absorptive properties of polymers. They design and test a non-electrical humidity sensor made of a polymer film.

Smart Sensors Module: Students investigate the behavior of pressure and heat-sensitive piezoelectric films. They use these materials to make a coin-counting machine and other smart sensing devices.

Sports Materials Module: Students explore the materials design and function of a wide variety of balls used in athletics. They also test and analyze the interaction of the balls with the many surfaces they come in contact with during play. Then students design a suitable material for use in a newly-invented game.

Department of Defense Sponsorship

As the largest federal employer of scientists and engineers, as well as a major R&D funder, the DoD has a core interest in assuring an ample national supply of U.S. technical talent. The National Research Council of the National Academy of Sciences and the National Academy of Engineering released a report on September 21, 2006 stating that kindergarten through eighth-grade science education is in "urgent need" of improvement. It also stated that science education is often based on flawed notions of how children learn and that "we are underestimating what young children are capable of as students of science."

DoD has identified *MWM* as a program that warrants support because it has the potential to spark interest in science and engineering while also equipping students with critical skills.

Four considerations have prompted DoD to make a multi-year commitment to scale up *MWM* in communities adjacent to DoD research laboratories and beyond: (1) field tests of *MWM* conducted by Northwestern University indicated effectiveness for all secondary students regardless of gender, income level, and race; (2) the program leverages a ready pool of DoD volunteer scientists and engineers across the country without making undue claims on their time; (3) *MWM* complements other DoD pre-college outreach programs that are delivered outside the classroom; (4) *MWM* has the potential to spur adoption in whole school districts and states.

DoD has piloted *MWM* over the past two years in middle and high schools near Picatinny, New Jersey; Aberdeen Proving Ground, Maryland; Indian Head, Maryland; and Benet Laboratory, New York. These pilot efforts have engaged almost 60 DoD scientists and engineers and more than 60 science teachers in partnerships to deliver *MWM*. The pilot phase enabled DoD to develop and refine the core components of a more far-reaching scale-up plan. This plan includes program introduction, teacher training, classroom implementation, and evaluation.

DoD has selected Maryland to begin a national rollout of *MWM* due to the state's concentration of DoD and other federal laboratories as well as its commitment to STEM education. Important groundwork has been laid in 2006 including the selection of Garrett College as the first DoD-sponsored State Center for STEM Education, the alignment of *MWM* to Maryland science standards, the piloting of a teacher training institute, the production of training videos, and the conclusion of a four-week experiment researching the impact of *MWM* on students' science knowledge and attitudes compared to a traditional teacher-directed curriculum. DoD has designated Action Research & Associates, Inc., to serve as outside evaluator of the program, to gather independent data on its effectiveness, independent of the field testing already conducted on *MWM* in various parts of the country.

Overall *MWM* Evaluation Plan

To answer the over arching question – "What is the impact of *MWM* on students' science learning, behavior, attitudes and aspirations?"— Action Research developed an "**Overall *MWM* Evaluation Plan.**" (See "Appendix A.") The objective of this Overall Plan is to gather sufficient data to produce scientifically-defensible and reliable information to answer this key question.

The three-phased "**Overall *MWM* Evaluation Plan**" to be implemented beginning 2007 is based on a quasi-experimental, mixed-method design, gathering both quantitative and qualitative data on a minimum of 2,000 students – 1,000 Treatment and 1,000 Comparison students. The data will be "triangulated." Multiple data sources will be assessed and compared to verify the reliability of the results. For instance, how *MWM* affects students' ability to problem solve will be assessed by student self-assessments, teacher and DoD volunteer assessments of the students, and classroom observations by trained researchers.

The data from the replicated studies over three years, 2007-2009, will provide a scientific, reliable and valid basis to guide DoD in making a well-reasoned and supported decision regarding whether or not to implement DoD-sponsored *Materials World Modules* programs in other locations in Maryland and throughout the Nation.

This report, which covers Phase I of the "**Overall *MWM* Evaluation Plan**," assesses the student and teacher outcomes of the 2006 Summer Institute funded by DoD at Garrett College.

Description Of The Study

Purpose of Study

As the first phase of a multi-year plan to measure the effectiveness of *Materials World Modules* (*MWM*), the Department of Defense (DoD) sponsored a Summer Institute from July 9-August 5, 2006, for 82 middle and high school students at Garrett College in Deep Creek Lake, Maryland. The Summer Institute was a residential immersion science and math institute designed to serve as the site for the first phase of a multi-year plan to measure the effectiveness of *Materials World Modules* (*MWM*)—an inquiry- and design-based materials science program—relative to a traditional teacher-directed curriculum.

DoD has made a multi-year commitment to support the scale-up of *MWM* in areas that host a significant concentration of DoD research facilities. A rigorous evaluation of *MWM* was deemed essential to the success of its statewide dissemination and the development of a national rollout.

Action Research & Associates, Inc.'s quasi-experimental research design included a two-tiered random selection process to 1) randomly select students to attend the Summer Institute; and 2) then randomly assign them to a Comparison or Treatment Group.

The Comparison Group experienced four weeks of traditional teacher-centered learning with science textbooks and labs. The Treatment Group received four weeks of student-centered learning, based on the *Materials World Modules* curriculum with labs and culminating design projects.

The following describes the program evaluation model for the *Materials World Modules*, and the components contributing to the overall design. This basic model was applied to evaluate the Phase I Summer 2006 Institute at Garrett College.

Using a **multi-method, quasi-experimental design with random assignment of students in matched pairs to treatment group**, the Summer Institute evaluation had **three main purposes**:

1. **To Gather Preliminary Evaluation Data on *MWM*:** The primary intent of the Phase I study was to initially explore the *MWM* Program's impact on students to assess whether further study of the *MWM* Program is merited. The study explored whether *MWM* improved students' content knowledge, ability to use scientific inquiry to problem solve, and aspirations to study science and pursue science careers. Employing a rigorous multi-method, quasi-experimental matched-pairs design, Phase I was designed to gather **preliminary** evaluation data to assess *MWM's* effectiveness.

2. **To Initiate Development of Reliable Assessment Instruments:** A second purpose of the Phase I study was to develop and pilot-test assessment instruments that can be used in a more comprehensive classroom evaluation in Maryland and elsewhere. If DoD decides to deploy the *MWM* Program throughout Maryland schools or in other states, a Formal Evaluation Study will be conducted in Maryland public schools during the regular school day and school year over several years. (See Appendix A for description of the Three Phase Overall Evaluation Plan.)

3. **To Provide Formative Feedback to Guide Training Teachers to Use *MWM*:** Finally, a third purpose was to analyze the Phase I data to provide DoD and other stakeholders with important formative information regarding what works and what refinements would strengthen the *MWM* Program and the associated Teacher Training Program. This would facilitate and guide their efforts in developing the next phases of *MWM* in Maryland and beyond.

Description of the Intervention

To provide an opportunity to conduct a rigorous preliminary and exploratory evaluation study, it was recommended that DoD sponsor a 2006 *MWM* Maryland Summer Institute. The Summer Institute focused on four science units - Composites, Sports Materials, Concrete and Polymers, with one unit taught each week. Every effort was taken to control key aspects of the students' environment during the four-week, 24-hour, seven-day week residential institute to ensure that only the independent variable of interest – *MWM* versus traditional classroom science methods and text – varied between Treatment and Comparison Groups.

- Treatment students received instruction using *MWM* materials, with student-centered instruction and labs, including iterative design projects and class presentations.
- Comparison students received instruction using traditional teacher-centered instruction supported by textbooks and teacher-directed lab experiences.

Other than the mode and materials of instruction, all students were treated in a similar way, from state-wide recruitment, two-tiered random selection to participate in the program and assignment to treatment group, to the size of their classes, quality of teachers, same number of hours of instruction, same tests and survey instruments, same science concepts taught in each module, same exposure to math and technology, and similar residential and social opportunities and activities.

Both Treatment and Control Group students received three hours of science instruction daily, in two 1.5 hour blocks, morning and afternoon, five days a week, for a total of 15 hours of science instruction weekly. They also both received three hours of math and technology (computers, calculators, etc.) instruction five days a week.

The two groups were generally separated during the school day (8am-5pm, Monday through Friday) and for many activities later in the day. The Treatment Group was taught in a separate building from the Comparison Group, and by different science teachers. They were taught by the same math teachers, with the Comparison Group being taught math in the morning and the Treatment in the afternoon. They ate lunch and dinner separately, and lived with roommates from their same group, generally on different floors in the residence hall. They enjoyed some social events, some limited free time and field trips together, but were transported on separate buses.

To ensure that the study was as "blind" as possible, the treatment groups were assigned coded names: Silver, Platinum and Nickel comprised the Treatment Group, and Copper, Gold and Iron comprised the Comparison Group. Upon arrival on campus, the students were directed to their assigned subgroup ("Copper" etc.) and they were not informed which treatment group they were in. In addition, the students were taught science in two different buildings and were physically separated during the day and both received high quality instruction. As a result, the students did not discover which treatment group they had been assigned to until the final week of the study. In addition, a weekly incentive program replaced the motivating factor of earning "grades" for academic work. Students in the Treatment and Comparison Groups received small-denomination gift certificates for achievement.

Prior to the institute, a curriculum committee developed common "Learning Objectives" for each unit, delineating key science principles to be covered in each. For instance, one Objective was to teach how to determine the "coefficient of restitution" or "COR" (how high a ball bounces). The same concept was taught in both groups, but how the concept was taught differed. The **Comparison Group** was taught science in the traditional manner, using predominantly teacher-centered lectures and lab experiences, and reading a science textbook (Pearson Prentice Hall's "*Physical Science: Concepts in Action: With Earth and Space Science*.")

The **Treatment Group** was taught with *Materials World Modules,* using a predominantly student-centered, inquiry- and design-based, hands-on approach. This involved *MWM* reading materials with *MWM* inquiry-based laboratory exercises and iterative project design experiences, and student presentations. Thus, the COR concept was taught to students using different materials and methods. The standardized pre-and post-module science tests were developed to measure knowledge of the Learning Objectives. (See Instrument Development.)

Description of the Teachers

Demographics of Teachers

A total of 11 individuals served as teachers in the Summer Institute (Table 2). In the Comparison Group, each of the three student groups (Copper, Gold, Iron) had the same teacher all four weeks of the program. In the Treatment group, the Nickel group had the same teacher all four weeks, the Silver group had four different teachers, and the Platinum group had three different teachers. This made a total of eight teachers in the Treatment group and three in the Comparison group. Each teacher completed a survey each week that elicited responses regarding his/her personal attributes, the module taught that week, and their perceptions of changes in the students.

Table 2 describes the demographic attributes of the *MWM* Summer Institute teachers. As a group, the Treatment teachers were older, ranging between 40 and over 60 years of age. In the Comparison Group, two teachers were under 40 while one was in the 50-59 age range. With respect to ethnicity and race, all teachers were non-Hispanic and white.

Table 2: Teachers' Demographic Attributes by Treatment Group

Attribute	*Treatment (N=8)*		*Comparison (N=3)*	
Age	*N*	%	*N*	%
20-29	0	0%	1	33.3%
30-39	0	0%	1	33.3%
40-49	2	25%	0	0%
50-59	4	50%	1	33.3%
60+	2	25%	0	0%
Gender				
Male	4	50%	2	66.7%
Female	4	50%	1	33.3%
Ethnicity				
Hispanic	0	0%	0	0%
Non-Hispanic	8	100%	3	100%
Race				
American Indian or Alaska Native	0	0%	0	0%
Asian	0	0%	0	0%
Black or African American	0	0%	0	0%
Native Hawaiian or Other Pacific Islander	0	0%	0	0%
White	8	100%	3	100%
Multiracial	0	0%	0	0%
Other	0	0%	0	0%

Professional Background of Teachers

Table 3 outlines the professional backgrounds of the Summer Institute teachers. Seventy-five percent of the Treatment teachers had 21 years or more of teaching experience. In contrast, two-thirds of the Comparison teachers had between 0 and 5 years of teaching experience.

The majority (65%) of the Treatment Group teachers were currently working as middle school or high school science teachers. The others were retired teachers. Among the three Comparison Group teachers, two were not currently working as teachers, but were working in educational institutions; the other was a certified high school science teacher.

In terms of educational background, the majority of the Treatment Group teachers (62.5%) had obtained a M.A. or M.S. degree; 25% had obtained a B.A. or B.S. degree, and one had earned a Master of Education degree. Among the Comparison group teachers, two (66.7%) held bachelor's degrees and one (33.3%) had earned a master's degree.

Table 3: Teachers' Professional Backgrounds by Treatment Group

Attribute	*Treatment (N=8)*		*Comparison (N=3)*	
Teaching Experience (Years)	***N***	**%**	***N***	**%**
0-5	0	0%	2	66.7%
6-10	1	12.5%	0	0%
11-15	0	0%	1	33.3%
16-20	1	12.5%	0	0%
21+	6	75%	0	0%
Science Teaching Experience (Years)				
0-5	0	0%	2	66.7%
6-10	1	12.5%	0	0%
11-15	0	0%	1	33.3%
16-20	1	12.5%	0	0%
21+	6	75%	0	0%
Current Teaching Position				
Middle school	1	12.5%	0	0%
Junior high school	0	0%	0	0%
High school	5	62.5%	1	33.3%
College	0	0%	0	0%
Other	2	25%	2	66.7%
Highest Degree				
A.A.	0	0%	0	0%
B.A./B.S.	2	25%	2	66.7%
M.A./M.S.	5	62.5%	1	33.3%
Ph.D.	0	0%	0	0%
Other	1	12.5%	0	0%

Description of Student Sample and Development Sample

Student Sample Development

The Summer Institute students were recruited from a pool of 401 middle and high school students in order to fill 96 spots. To develop this pool, in April 2006, Garrett College sent highly attractive, professionally-designed program announcements and posters to the Maryland Science Teachers Association, and to every Maryland school district via superintendents, principals, science curriculum coordinators, and science department chairs. Posters were included to alert students of the opportunity to apply to attend the Institute. Students were invited to learn more and apply for the Institute via the Internet.

All Maryland secondary students were eligible and encouraged to apply. To encourage as broad a sample as possible, there were no screening criteria regarding minimum grade point average or number or types of science courses studied. In addition, financial cost was not a barrier because DoD paid the expenses for all selected students, enabling students from a wide range of socio-economic backgrounds to attend. The stipend covered their room and board, travel expenses, recreational activities, and stipend. Students came from virtually every county across Maryland from large inner city schools, private schools, parochial schools, charter schools, large public schools in suburban and urban areas, military academies, technical schools, and "home schools." Based on the students' self-reported motivations for attending, it is clear that some students applied primarily motivated by the science immersion experience, and others by the stipend and monetary incentives. The above combined factors led to a broader pool and sample of secondary students than would normally occur for a summer gifted and talented program or traditional science camp. As a result, it was possible to recruit and select a diverse cross-section of students from every part of the state.

The applicant pool resulted in a two-to-one ratio, middle school (261) to high school applicants (140). (See Table 4) For the purposes of this program, "middle school" was defined as grades 6-9 and "high school" was defined as grades 10-12. According to Garrett College officials, this disparity in applicants' grade levels is not surprising. In their experience administering similar summer science programs for more than 15 years, middle school students tend to be more motivated to attend organized summer programs, their parents are more involved in making their summer plans, and they are less likely to seek and find summer employment than high school students. As a result, it was determined that each treatment group would consist of one high school class and two middle school classes, of approximately 15 students each.

Sample Selection and Assignment to Treatment Group

To select the final sample, a two-tiered stratified random sampling process was used to create a probability sample to: 1) select the attendees from a pool of 401 applicants, and match them by demographics (grade, race, sex), and 2) assign the participants to a treatment group. A two-tiered stratified random selection process reduced the chance of bias potentially influencing study results. The probability sample also allowed the researcher to make statistical inferences about the population, based on the results obtained from the sample. Per Alreck and Settle:

> "With random sampling, the researcher can calculate and report the "statistical significance" of relationships... based on the probability that such relationships would result only from sampling error." (p. 69**,** *The Survey Research Handbook*. Richard D. Irwin, Inc. Homewood, IL, © 1985.)

A sampling model was developed by Action Research & Associates, Inc., the independent research firm conducting the evaluation under contract to BEST. The sampling model reflected Maryland's demographics regarding race and sex, as well as the proportion of students' grade-levels, adjusted proportionately to a sample size of 96 participants. Applicants were stratified by grade, race and sex, and randomly selected by matched pairs to fit the demographic proportion in the sample model. Next, the matched pair participants were randomly assigned to either the Treatment or Comparison Group. (Table 4)

Table 4: Sampling Model

Applicants			Selected Students						
Grade Level	# of Applicants	% of Applicants (N=401)	# of Selected Students	% of Selected Students (N=96)	Sex	Race: Black	Race: White	Race: Other	Total
7th	59	15%	20	20.8%	Female	4	6	0	8
					Male	2	6	2	10
					Total	6	12	2	20
8th	100	25%	22	22.9%	Female	2	8	0	10
					Male	4	6	2	12
					Total	6	14	2	22
9th	102	25%	22	22.9%	Female	4	6	2	12
					Male	2	8	0	10
					Total	6	14	2	32
Middle School Subtotal	*261*	*65%*	*64*	*66.7%*	*Female*	*10*	*20*	*2*	*32*
					Male	*8*	*20*	*4*	*32*
					Total	*18*	*40*	*6*	*64*
10th	58	15%	16	16.7%	Female	2	4	2	8
					Male	4	4	0	8
					Total	6	8	2	16
11th	55	14%	12	12.5%	Female	2	4	0	6
					Male	2	2	2	6
					Total	4	6	2	12
12th	27	6%	4	4%	Female	2	0	0	2
					Male	2	0	0	2
					Total	4	0	0	4
High School Subtotal	*140*	*35%*	*32*	*33.3%*	*Female*	*6*	*8*	*2*	*16*
					Male	*8*	*6*	*2*	*16*
					Total	*14*	*14*	*4*	*32*
TOTAL	**401**	**100%**	**96**	**100.0%**	**Female**	**16**	**28**	**4**	**48**
					Male	**16**	**26**	**6**	**48**
					Total	**32**	**54**	**10**	**96**
					%	33.3%	56.3%	10.4%	100%

Description of Final Sample

In the final sample, of the 82 who completed the four-week Institute, the proportion of Treatment and Comparison students was the same – 41 each. (Table 5)

By Grade Level

Both treatment groups had 13 high school and 28 middle school students each. Both groups consisted of one high school group of 13 students, and two middle school groups, with 14 students each. (Table 5)

Table 5: Sample by Grade-Level and Treatment Group

Grade	*Sampling Model Goal (%)*	*Total*	*% of N (82)*	*Treatment*	*Comparison*
6th	—	3	3.7	0	3
7th	15	14	17.1	8	6
8th	25	25	30.5	13	12
9th	25	14	17.1	7	7
Middle School Subtotal	**65**	**56**	**68.4**	**28**	**28**
10th	15	11	13.4	6	5
11th	14	14	17.1	7	7
12th	6	1	1.1	0	1
High School Subtotal	**35**	**26**	**31.6**	**13**	**13**
TOTAL	**100%**	**82**	**100.0%**	**41**	**41**

By Race

Proportions of Black, White, and "Other" students in the final sample were 32.9%, 59.8%, and 7.3%, respectively (Table 6). This is comparable to the U.S. Census Bureau's statistics for Maryland, which lists percentages of 28.7% Black, 61.5% White, and 9.8% "Other." Every effort was made to split each racial category evenly between the Treatment and Comparison Groups.

Table 6: Sample by Race and Treatment Group

Race	*Maryland Racial Composition (%)*	*Total (N)*	*% of N (82)*	*Treatment*	*Comparison*
Black	28.7	27	32.9	12	15
White	61.5	49	59.8	26	23
Other	9.8	6	7.3	3	3
TOTAL	**100%**	**82**	**100.0%**	**41**	**41**

By Gender

In the final sample, the proportion of male and female students was evenly split, with 41 of each in both treatment groups. (Table 7) The even split was maintained across middle school and high school sub-samples. Again, this is comparable to the U.S. Census Bureau's statistics for the State of Maryland, in which 48.2% of residents are male and 51.8% are female. Equal numbers of male and female students (+/- 1) were assigned to the Treatment and Comparison Groups.

Table 7: Sample by Gender and Treatment Group

Gender	*Maryland Gender Composition (%)*	*Total (N)*	*% of N (82)*	*Treatment*	*Comparison*
Male	48.2	41	50.0	21	20
Female	51.8	41	50.0	20	21
TOTAL	**100%**	**82**	**100.0%**	**41**	**41**

Attrition

Through the course of the Institute, a total of 18 students were lost due to attrition, (Table 8) due to first day "no-shows," homesickness, discipline issues and family vacations. Twelve students selected to participate failed to attend the first day and were replaced by additional first-day recruitment of four local students who had applied, who could respond quickly to fill in.

Analyzing the data, high school and middle school students were equally likely to fail to show the first day of the Institute. However, almost all of the attrition (83%) during the Institute occurred among the high school students. There were negligible differences in attrition by sex or race.

Table 8: Sample Attrition

Reason	Sex	Black	White	Other	Grade		Comparison/ Treatment	Total Loss/Gain
Pre-Assignment Losses								
Selected, No-Shows (Failed to Attend on First Day)	Male: 7 Female: 5	3	5	4	7th 8th 9th 10th 12th	2 1 3 5 1	Not Applicable	-12
Post-Assignment Losses								
Homesick (Left During First Week)	Male: 1 Female: 2		2	1	9th 10th 11th	1 1 1	Comparison: 2 Treatment: 1	-3
Family Vacation	Male: 0 Female: 1	1			11th	1	Comparison: 1	-1
Disciplinary	Male: 2 Female: 0	2			10th 12th	1 1	Treatment: 2	-2
Total Leaving Program (Post-Assignment)	Male: 3 Female: 3	3	2	1	9th 10th 11th 12th	1 2 2 1	Comparison: 3 Treatment: 3	-6
Post-Assignment Gains								
Students Recruited to Replace First Day No-Shows	Male: 2 Female: 2		3	1	6th 11th	3 1	Comparison: 4	+4

Looking at the percentage of students on medications other than allergy or cold medicines, about 60% of the students did not use any medication. Approximately 40% of the final sample did regularly use medications. Documented medications being used by this sub-sample of students reflected treatment of a variety of conditions, including ADHD and other learning and behavioral disorders including social-emotional disorders.

Research Model

To assess the effectiveness of the *MWM* Program with secondary school science students, the Phase I study developed a series of questions to be investigated empirically. These questions form the heart of the research plan and were developed to guide the research work. The research questions reflect the objectives of the evaluation and provide a practical structure and a framework for stakeholders.[1]

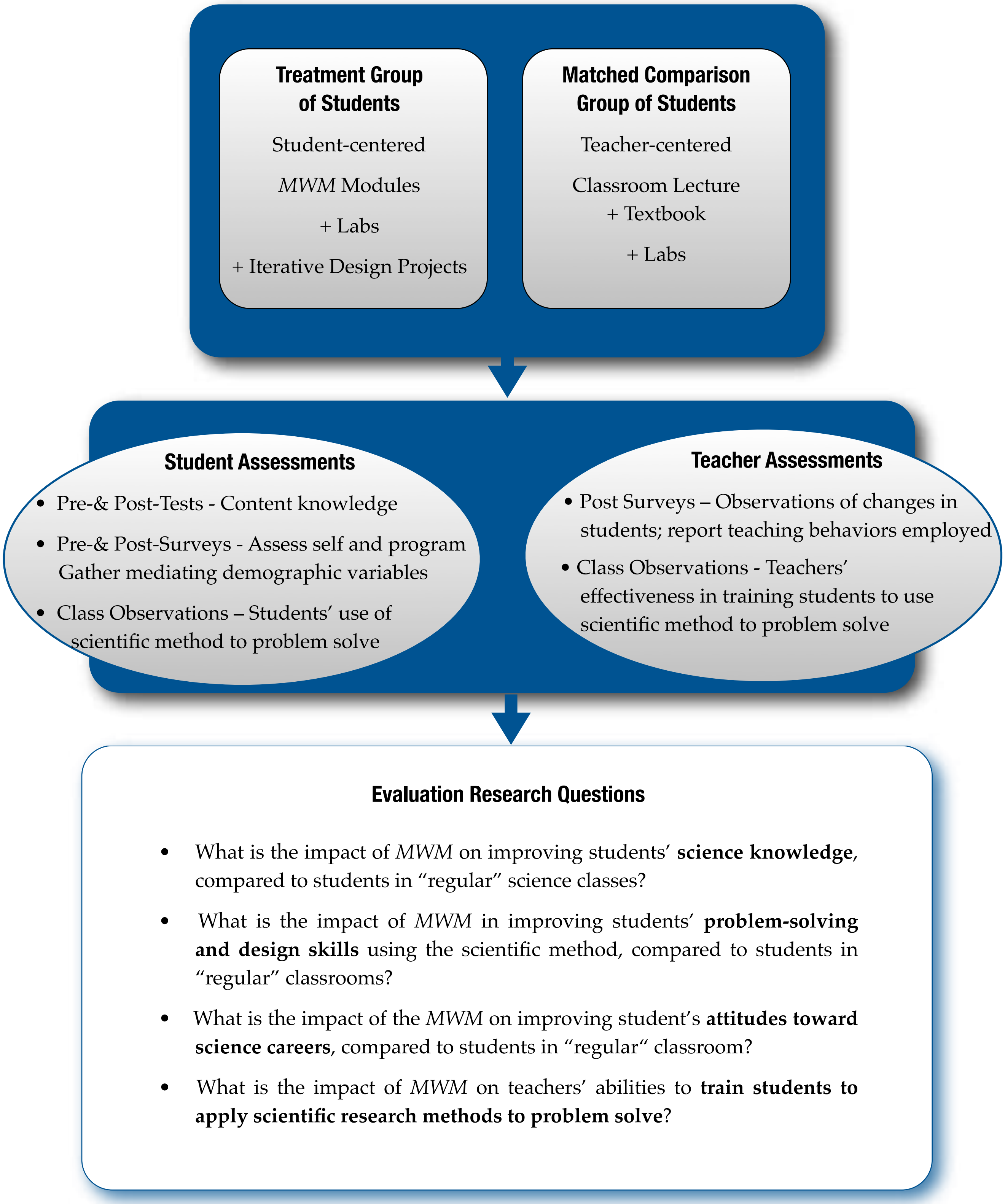
Research Questions

1. Compared to students in "regular science classes" with teacher-led, textbook instruction (Comparison Group), what is the impact of *MWM* Modules (Treatment Group) on improving students' science knowledge?
2. Compared to students' in regular science classes, what is the impact of the *MWM* Modules on improving students' use of the scientific method to problem solve and design projects?
3. Compared to students in regular science classes, what is the impact of the *MWM* Modules on improving students' attitudes toward the study of science or careers in the science field?
4. Compared to teachers in regular science classes, what is the impact of the *MWM* Modules on improving science teachers' abilities to train students to apply scientific research methods to problem solve?

The figure on the following page illustrates the research model and related research questions employed in the Phase I *MWM* evaluation (Figure 3):

[1] **Scientific Research in Education** edited by Shavelson, R.J. and Towne, L. Committee of Scientific Principles for Education Research, Center for Education, Division of Behavioral and Social Sciences and Education, National Research Council, National Academy Press, Washington, DC, 2002.

Figure 3:
Evaluation Model for *Materials World Modules*



Data Collection Methods and Instruments

Overview

A series of key factors were derived by analyzing the four research questions. The key factors to be measured included:

1) **Changes in Students'**

   a) science content knowledge;

   b) ability to apply scientific inquiry to problem solve;

   c) interest in further science study;

   d) career aspirations;

   e) science skills; and

   f) team behaviors.

2) **Teachers' perceptions of**

   a) students' cognitive gains;

   b) students' ability to use scientific inquiry to problem solve; and

   c) their own ability to 1) apply scientific inquiry methods in teaching and 2) coaching students to problem solve in classroom, laboratory and project-design activities.

To answer the four proposed research questions and measure relevant changes, a series of scientifically-developed instruments were required to be developed: Action Research & Associates, Inc, an independent research firm with test development expertise, developed the series of data collection instruments measuring cognitive and attitudinal changes in students. The student data were triangulated with their teachers' perceptions of the students' cognitive, attitudinal and skill gains as well as classroom observations by the independent researcher. The following data gathering instruments were developed:

Student-level data was gathered via the following instruments:

1) **Pre- and Post-Module Tests:** Weekly tests were administered before and after each module was taught to measure student content knowledge gains.

2) **Student Application Form**: Basic student demographic data were gathered prior to admission, including grade level, educational background, city and county, and student and parent contact information.

3) **Pre- and Post-Institute Student Surveys:** Surveys were administered before and after the four-week Institute to capture changes in students' self-assessed

   a. interest in studying science and career aspirations;

   b. ability to use the scientific method;

   c. ability to problem solve;

   d. team work skills; and

   e. science skills.

4) **Classroom Observations:** To assess how Treatment and Comparison students:

a. use the scientific inquiry method in problem solving in labs and projects;

b. apply their newly gained content knowledge;

c. collaborate in a team setting; and

d. use the precise language of science, mathematics and engineering.

Teacher-level data was gathered via:

Teacher Post-Module Surveys: Information from Treatment and Comparison Group teachers instructing in the four-week Summer Institute was gathered every week after completing the module. This was necessary because the duration a teacher taught was expected to vary from one week to multiple weeks.

The Teacher Post-Institute Survey gathered weekly data on teachers':

a. demographic and education backgrounds;

b. self-assessed roles in the classroom, the teaching process, and Comparison and Treatment materials taught;

c. self-evaluation of their ability to better use scientific inquiry in the classroom to teach and coach students to use scientific inquiry to problem solve, and

d. perceptions of students' changes in attitudes toward science, science and inquiry skills and ability to problem solve in the classroom.

Research Questions and Sources of Data

Table 9 summarizes the four research questions and the data sources to address them:

Table 9: Research Questions and Sources of Data

Comparing the Treatment Group with the Comparison Group, what is the impact of MWM on improving...	*Data Sources*				
	Teachers		*Students*		
	Post-Institute Survey	*Class Observations*	*Pre- Post-Institute Surveys*	*Pre- Post-Module Tests*	*Class Observations*
1 …students' science knowledge?	✓		✓	✓	✓
2. …students' problem solving and design skills using of the scientific method?	✓	✓	✓	✓	✓
3. …students' attitudes toward studying science and careers?	✓		✓		
4. …science teachers' abilities to train students to apply scientific research methods?	✓	✓	✓	✓	✓

Development and Administration of Instruments

Student-Level Instruments

Pre- and Post-Module Content Knowledge Tests

The Standardized Pre- and Post-Tests for the Modules administered to both Treatment and Control Group students to measure knowledge gains were developed employing the following steps:

1) *MWM* Summer Institute Program Developers reviewed the nine *MWM* modules and selected four with concepts and activities that were appropriate and adaptable for middle school, as well as high school students. Four *MWM* modules were selected, one for each week of the Summer Institute: *Composites, Sports Materials, Concrete,* and *Polymers.*

2) The Summer Institute Program Developers identified and listed the key scientific concepts and supportive math concepts included in each of the four *MWM* Modules. The developers identified between four and eight" Learning Standards" or instructional objectives for each Module to bridge the content of *MWM* and the selected science textbook. This included concepts such as the coefficient of restitution, coefficient of friction, compression and tension, torque and elasticity, strength and stiffness, properties of molecules, phases of matter, kinetic and potential energy, etc.

3) The Institute developers surveyed conventional textbooks currently used by Maryland public schools to identify the text and chapters that best covered these key concepts. They identified a primary text to use with the Comparison Group, *"Physical Science: Concepts in Action: With Earth and Space Science"* published by Pearson Prentice Hall, Upper Saddle River, NJ.

4) Potential test items for field testing were reviewed from a variety of public sources such as science items posted on the Internet from the National Assessment of Educational Progress (NAEP), American College Testing (ACT), and items from state assessments from Maryland, Virginia, Texas, Ohio, Florida, and Delaware, and other tests.

5) To correlate the items as closely as possible with the content of the four modules for the Treatment and Comparison Groups. The following sources were primarily used for constructing the field tests of the four module tests:

   a) ***MWM* Test Item Bank**: The test item bank from which this study selectively adopted multiple-choice and rubric essay questions was originally developed by a panel of science educators working with *MWM* at Northwestern University with funding from the National Science Foundation. The panel produced assessment items for a series of on-line *MWM* modules, known as *MWM 2002*, intended primarily for high school students. The objective was to create a bank from which teachers could select and customize their own tests to match their classroom goals. Because many of the activities in the *MWM 2002* series were similar or identical to the *MWM* (booklet version) used in the Summer Institute, and because the items had undergone content validation, the bank of *MWM 2002* items provided a logical starting point. The items that accompanied each *MWM* module had not been previously analyzed for their psychometric properties, nor were they validated or standardized for use in a research study. The 2006 Summer Institute provided the first opportunity to pilot test the items to empirically explore their usefulness in terms of reliability to create new standardized tests.

b) ***"Chapter and Unit Tests: Levels A and B"*** - test items were selected from the Pearson Prentice Hall science textbook used for the Comparison Group. No psychometric data on the items was provided by Pearson Prentice Hall.

6) Based on the *MWM* field testing, several of the Prentice Hall and the *MWM* Test Item Bank items used in the *MWM* field tests proved to be flawed and were eliminated based on empirical statistical data. From these sources, a pool of field test items was developed to reflect each Module's content and "Learning Standards."

7) The final items were selected to be field tested for each module to reflect the following:

   a) the range of Learning Standards;

   b) a variety of question formats – multiple choice, short and long answer rubric questions; and

   c) multiple levels of Bloom's ***Taxonomy of Educational Objectives***[2] including Knowledge, Comprehension, Application, Analysis, and Synthesis.

8) Approximately half of the items on each test were from the *MWM* Test Item Bank and half from the Pearson Prentice Hall textbook test items.

9) If more than one test item relating to a unit Learning Standard was identified, a decision was made as to which item was most closely aligned with the standard. Occasionally, decisions were made to include more than one Learning Standard-related item on the test to empirically explore how the item functioned in the test environment.

10) Test items were reviewed and edited by a team consisting of a university and school science expert and a DoD test and survey expert for content validity, clarity, ambiguity and reading level.

11) The same four tests of content knowledge (Composites, Sports Materials, Concrete, Polymers) were administered to both Treatment and Comparison Groups, with the pre-test administered the day before starting the module (usually Sunday) and the post-test at the end of study (usually Friday). Subsequently, revised tests were developed based on psychometric and content analyses thereby creating the standardized tests.

Student Application Forms

The Student Application gathered basic student demographic data. This formed the basis for creating a stratified sample pool of students by grade, sex and race, from which the students were randomly selected for the 2006 *MWM* Summer Institute as well as for assignment to the Treatment or Comparison Groups.

Pre- and Post-Institute Student Surveys

The Student Surveys were administered before and after the four-week Institute. These instruments consisted of two parts.

a) ***Part One-Science Careers*****:** Assessed and captured changes in students' perception of the desirability of science careers, and their aspirations, as well as their perceptions of the study of science. This was assessed primarily via five-point Likert scales and open-ended questions.

[2] From Benjamin S. Bloom, ***Taxonomy of educational objectives***. Published by Allyn and Bacon, Boston, MA. Copyright © by Pearson Education.

b) ***Part Two-Student Self-Assessment***: Provided feedback from the students' perspective regarding self-assessed changes in their knowledge and ability to use the scientific method to problem solve. Surveys were developed to assess and capture the students' experiences with their school year instruction versus the instruction at the Summer Institute and to determine if there were any differences between Treatment and Comparison Groups in students' self-perceived growth in teamwork skills attributable to the Summer Institute experiences.

Classroom Observations

The purpose of this form was to capture the use of scientific inquiry in the classroom based on teacher-to-student and student-to–student interactions as the students worked in their labs and design projects. This protocol was developed based on Horizon Research, Inc.'s ***Classroom Observation Form***,[3] and the ***5 E Model*** (Perkins, 1992).[4] Additional elements were incorporated to accommodate the *MWM* project design elements.

Teacher-Level Instruments

Teacher Surveys

Teacher surveys were administered at the end of the Summer Institute to measure teachers':

a) self-evaluation of changes in their ability to use scientific inquiry in teaching and coaching students.

b) perceptions of students' changes in teamwork skills within cohort groups attributable to the Summer Institute program.

Classroom Observations

Classroom observations assessed how teachers instructed the Treatment vs. Comparison Group students, and the degree to which they coached students to use the scientific method to problem solve and design their projects. This protocol was developed based on Horizon Research, Inc's Classroom Observation Form, and the "5 E Model" (Perkins, 1992).[4] Additional elements were incorporated to accommodate the *MWM* project design elements. Both Comparison and Treatment Teachers were observed by a trained independent educational researcher a minimum of two times during the Institute.

[3] Weiss, I. et al, ***Looking inside the classroom: A study of mathematics and science education in the United States.*** Chapel Hill, NC Horizon Research, Inc.: May 2003.

[4] Perkins, D. 1992. ***Smart schools: Better thinking and learning for every child.*** New York: The Free Press, as reproduced in the National Institutes of Health Curriculum Supplement Series: Grades 9-12, ***Emerging and Re-emerging and Infectious Diseases***. Copyright © 1999, BSCS and Videodiscovery, Inc.

Table 10 summarizes the data instruments and the schedule for administering them.

Table 10: Data Instruments and the Schedule of Administration

Assessment Instruments Administered	*June 1-17*	*June 18-24*	*June 25 - July 1*	*July 2-8*	*July 9-15*	*July 16-22*	*July 23- 29*	*July 30- Aug 5*
Teachers								
Post Institute Survey					X	X	X	X
Class Observations					X		X	X
Students								
Application Form	X							
Pre-Institute Survey					X			
Post-Institute Survey								X
Pre-Test/ rubric questions					X	X	X	X
Post-Test/ rubric questions					X	X	X	X
Class Observations					X		X	X
Focus Groups						X		
Mid-Institute Survey						X		

Description Of Data And Statistical Analysis

Pre- and Post-Tests

For the purposes of the Phase I *MWM* Evaluation Study, the field test items on the knowledge tests and surveys were analyzed using item analyses and other psychometric analyses, and their correlation and reliability estimates relative to the overall score, scales and subscales were computed. Those items empirically determined to be of low reliability, ambiguous, or that did not differentiate well were eliminated and a revised knowledge test was developed, *ex post facto,* thereby producing a revised test with increased reliability and content validity. Reliability estimates and other statistics were developed for the revised tests (see Table 11). Each module test ranged between 20 and 33 items with approximately half of the items coming from *MWM* and half from Pearson-Prentice-Hall's test items. The module tests were aggregated to create an overall test score.

Table 11: *MWM* Module Tests' Reliability Estimates

MWM Module Tests	*Number of Items*	*Number of Valid cases*	*Reliability Estimates**
Composites	25	84	.75
Sports Materials	33	88	.78
Concrete	20	84	.68
Polymers	25	80	.84
Combined tests	93	75	.93

* *Cronbach's coefficient alpha*

- **Students' growth in content knowledge** was assessed by comparing the difference between the pre- and post-test scores for the modules for the Treatment and Comparison Groups. Treatment Group students were matched by grade level with Comparison Group students. Aggregate group scores were statistically compared to measure growth across and between Treatment and Comparison cohort groups. This facilitated examination of whether *MWM* exposure produced a different and significantly positive gain in content knowledge.
- The student scores were computed and the **students' growth in knowledge**, pre- to post-test, was calculated and averaged. The Percent Value Added, Standardized Mean Gain Effect and the Normalized Gain were calculated (see below). The results were analyzed by sex, race and grade level, and compared between the Treatment and Comparison Groups. Follow-up t-tests were calculated to determine the level of significance.

Students' and Teachers' Surveys

- Statistical analyses were conducted to determine whether the Treatment or Comparison Groups exhibited a **change in their ability to apply new-found knowledge and science skills** in a meaningful way **in the labs and project designs**. Areas of evaluation included, but were not limited to, gains in the students' ability to 1) use scientific inquiry, 2) critically think and problem solve, 3) collaborate in a team setting, 4) use the precision of the language of science, mathematics and engineering, 5) apply the scientific method to identify, formulate and execute a laboratory experiment or design project using application of concepts, analysis and problem-solving skills, and 6) gains in the number of students aspiring to careers in the science field.
- Data were analyzed to assess whether students in the Treatment or Comparison Groups have similar or different **changes in attitudes regarding the study of science in general, and specifically careers** in science. The changes in students' attitudes, behaviors, science skills, team behavior pre – to post-Institute, gathered via 1) pre- and post Institute surveys, were calculated by computing the mean percent change with follow up t-tests to determine the level of significance of changes between the Treatment and Comparison Groups.
- Quantitative and qualitative analyses of **teachers' and students' assessments of their *MWM* experience** were conducted.

Statistical Analysis

A number of statistical indicators are commonly used by educational researchers to evaluate the effectiveness of a teaching intervention. Each of the statistical indicators has both benefits and drawbacks. As a result, three indices of science knowledge gain are used in this evaluation. These are listed below in order of their ease of use.

1. Percent Value Added

The "Percent Value Added" (or Simple Percent Gain) is useful as an indicator because it is relatively easy to understand.

The equation for computing Percent Value Added (PVA) for a group is:

PVA= group average post-test percent score – group average pre-test percent score

For example, a PVA of 42% means that the group has increased its knowledge by 42% more than they knew before the intervention.

2. Normalized Gain <g>

Another important indicator used in educational program evaluations is the "Normalized Gain" or <g>. The <g> statistic is normalized by computing the average percentages of pre-test and post-test items answered correctly:

A Normalized Gain of .25 means that the group, on average, gained the equivalent of 25% of the maximum gain possible on the test. Another way of viewing this is that the group progressed 25% beyond the average pre-test score towards the maximum score on the test. This statistic has been used to evaluate the teaching of undergraduate courses in physics and engineering using a hands-on, inquiry-based approach. Hake

(1998a, 1998b)[5] recommends the following criteria for interpreting study results:

<g> 0 - .30 = small gain

<g> .31 - .70 = moderate gain

<g> .71 – 1.0 = large gain

3. **Standardized Mean Gain Effect**

The "Standardized Mean Gain Effect" size is an indicator of knowledge gain widely used in educational research. The Standardized Mean Gain Effect is appropriate to apply when a study is designed to compare two dependent groups, [6] such as matched pairs, like the *MWM* study. Effect size (ES) is a name given to statistics that measure the magnitude of a treatment effect. The generally accepted standards for educational and psychological testing (1999) [7] recommend that effect size and its corresponding 95% confidence interval (CI) be reported as part of the study results. The standardized mean gain effect size can be compared across groups and used in meta-analyses. Effect size reports the magnitude of a treatment effect as measured in standard deviation units. For example, an effect size of .72 means that a group has improved .72 standard deviation units from a mean pretest score to a mean post test score.

---

[5] Hake, R.R. 1998a. "Interactive-engagement vs. traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses," **Am.J. Phys** 66:64-74; online as ref 24 at http://www.physics.indiana.edu/hake./

Hake, R.R. 1998b. "Interactive-engagement methods in introductory mechanics courses." Online as ref. 25 at http://www.physics.indiana.edu/hake./

[6] Lipsey, M.W. & Wilson, D.B., 2001. ***Practical meta-analysis.*** pp. 44-46. Thousand Oaks, CA. Sage Publications.

Becker, B.J., 1988. "Synthesizing standardized mean-change measures." ***British Journal of Mathematical and Statistical Psychology,*** Vol. 41, pp. 257-278.

[7] ***The Standards for Educational & Psychological Testing***, developed jointly by the American Education Research Association (AERA); the American Psychological Association (APA; and the National Council on Measurement in Education. AERA Publication Sales: Washington, DC,: 1999.

Institutional Review Board

Due to the design of this evaluation, which involves assessing children ages eleven through nineteen in an academic setting and collecting demographic and other information from the parents of these children and their teachers, it was necessary to apply for Institutional Review Board (IRB) approval before proceeding with data collection. Action Research & Associates, Inc. worked with the Walter Reed Army Institute of Research (WRAIR) to apply for IRB approval from the Human Use Review Committee of the Walter Reed Army Institute of Research by submitting a modification to the previously-approved protocol WRAIR HURC Protocol Number 1216.

As part of the IRB approval process, all key participants in the *MWM* Summer Institute, including all researchers, and Summer Institute program directors, supervisors and teachers, took a National Institutes of Health on-line course and completed certification in "Human Participant Protections Education for Research Teams." The WRAIR Human Use Review Committee provided unbiased consultation that ensured the study maintained the legal rights of all participants and the work met the highest ethical standards for research. WRAIR reviewed and approved the study design, instrumentation, and research plan in May 2006 prior to the research team entering the field.

Results

Introduction and Study Objectives

The 2006 Summer Institute, a small-scale study, provided an opportunity to preliminarily evaluate the effectiveness of *MWM* in terms of gains in 1) students' science knowledge, 2) science-related attitudes and skills relative to their counterparts receiving traditional teacher-led science instruction, textbooks and labs, as well as 3) effects of teacher behaviors on students. The "Results" section is segmented to present the most relevant data.

How to Interpret the Results

In interpreting these results, it is important to note that this study is the first phase of a multi-phased evaluation, and as such, it was designed to develop *preliminary and exploratory data* to assess the potential merits of *MWM* relative to traditional classroom science instruction. This was a *small-scale pilot study in preparation for a full-scale evaluation.* The full-scale evaluation with a sample of an estimated 2000 students, is planned for 2007-2009, will provide the statistical power to be able to project the findings to the general population of Maryland students.

To understand these results, keep in mind that:

> "Tests of statistical significance compare groups of data to determine the probability that differences between them are based on chance, thereby providing evidence for judging the validity of a hypothesis or inference.…When a difference between two means is significant at the $p<0.05$ level, this means that the probability is less than 5 out of 100 times that the difference is due to chance. On this basis, it is possible to conclude that the differences obtained were the result of the treatment." pp 249-250, Tuckman, B. *Conducting Educational Research, Second Edition.* © Harcourt, Brace, Jovanovich, Inc., 1978.

In addition, in interpreting the results, it is important to remember the following about the Comparison and Treatment samples:

- The students were not pre-screened on the basis of academic merit.
- They were randomly selected from a pool of 401 applicants from throughout the State of Maryland to be representative of the state's demographics.
- A sizeable number of the students were from inner city schools, were minority or from other disadvantaged groups.
- Two-thirds of the students were middle school students going into 6^{th} - 9th grades.
- The *MWM* modules were originally designed for high school students, and were adapted by the Summer Institute teachers for middle school students.
- The *MWM* test items were originally developed for high school students, and were not modified for the Institute middle school students.
- Twenty-two percent of the students were taking medications for learning and behavioral disorders during the Summer Institute. These students were generally evenly distributed between the two groups.

- Whereas many of the students self-selected to attend the Institute because of interest in science, many of the students stated in the Pre-Institute Survey they had applied primarily for non-academic reasons such as to obtain the stipend.
- A number of statistical indicators are used by educational researchers to evaluate the effectiveness of a teaching intervention. Each of the statistical indicators has both benefits and drawbacks. As a result, several indices of program performance are used in this report. To more fully understand the differences between these statistical measures of gain, refer to pages 33-34, "Statistical Analysis."

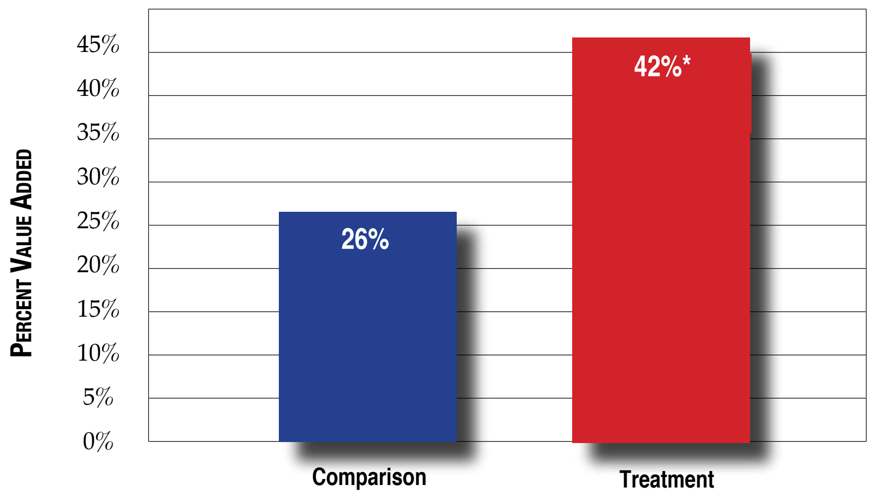
The evaluation objective was to assess the *potential* benefits of *MWM* **relative to traditional classroom instruction, to prepare for a full-scale evaluation that will assess additional variables. With this qualification in mind, the following *preliminary findings* stand out:**

MWM's Impact on Students' Knowledge of Science

Overall Knowledge Gains - Comparison vs. Treatment Students

- Due to the two-tiered randomization process used to 1) select the sample in matched pairs and 2) to assign students to Treatment and Comparison Groups, **both Groups started the Institute with the same level of science knowledge** relative to the units of study. That is, the two Groups' average scores were the same overall on the pre-tests.
- To determine the value added by learning science through *MWM* Modules relative to traditional textbook-centered approaches, the value added percent was calculated for Treatment Groups and by Key Demographic Characteristics. (See the "Statistical Analysis" Section for details.) Based on those calculations, the following was found:
- The Treatment students – those taught with *MWM*—made statistically significant ($p<0.0001$) knowledge gains over their Comparison Group peers after the same number of hours of instruction. That is to say, **it is very highly likely that the differences in pre- to post-test scores between the two Groups was due to the differences in treatment — the *MWM* Modules**.
- After studying the *MWM* units, the Treatment students significantly ($p<0.0001$) improved their pre- to post-test science knowledge scores by an average 42% gain relative to the Comparison Group students with an average gain of 26%. (Figure 4, Table 12)
- This would suggest that **for the same 15-hour time investment, students learning via *MWM* made significantly larger gains in science knowledge over the Comparison Group** using traditional classroom materials and methods.

Figure 4
Students' Overall Gains in Science Knowledge - Percent Value Added



Base: Pre-Test Scores
* statistically significant at p<0.0001

Table 12: Students' Overall Gains in Science Knowledge – Percent Value Added

Module	*Group*	*Percent Value Added*	*t-test*
Total	Comparison	26%	p<0.0001
	Treatment	42%*	

** significant at p<0.05*

Students' Gains in Science Knowledge by Module — Percent Value Added (PVA)

After both groups received 15 hours of instruction to cover the Learning Objectives in each module:

- The Treatment Group significantly out performed the Comparison Group on the Composites test, (35% to 23% PVA), and the Polymers test, (52% to 16% gain). (Table 13)
- The Treatment Group out performed the Comparison Group on the Sports Materials test (28% vs. 17% PVA), which approaches the level of significance ($p<0.08$).
- Only on the Concrete Module did the Comparison Group outperform the Treatment Group (59% to 51% PVA, not statistically significant). This reversal in trends in gains is most likely linked to the Comparison Group's deviation from the standard test preparation protocol that week. Both Groups fully complied for the other three modules.

Table 13: Students' Gains in Science Knowledge by Module - Percent Value Added

Module	*Group*	*Percent Value Added*	*t-test*
Composites	Comparison	23%	p<0.045
	Treatment	35%*	
Sports Materials	Comparison	17%	p<0.08
	Treatment	28%	
Concrete	Comparison	59%	p<0.38
	Treatment	51%	
Polymers	Comparison	16%	p<0.0001
	Treatment	52%*	

** significant at p<0.05*

Students' Gains in Science Knowledge — Normalized Gain

- To determine the value added by learning science through *MWM* Modules relative to traditional textbook-centered approaches, Normalized Gain was computed for both groups. The Comparison and Treatment Groups had been randomized so effectively that both Comparison and Treatment students had identical average pre-test scores overall – they knew 33% of the material as measured by the pre-tests. (Table 14)
- After both Groups received 15 hours of instruction weekly, the Group taught by *MWM* showed significantly ($p<0.01$) higher normalized gains (49%) overall, compared to the Comparison Group (45%).
- **This would preliminarily indicate that for the same time investment, *MWM*-trained students made significantly larger gains in science knowledge over the Comparison Group.**

Table 14: Students' Overall Gains in Science Knowledge
Normalized Gain

Module	*Group*	*Normalized Gain*	*t-test*	*Pre-test*	*Post-test*
Combined Tests	Comparison	0.18		33%	45%
	Treatment	0.25*	0.01	33%	49%

* *significant at p<0.05*

Table 15: Students' Overall Gains on Science Knowledge by Module
Normalized Gain

Module	*Group*	*Normalized Gain*	*t-test*	*Pre-test*	*Post-test*
Composites	Comparison	0.14		32%	42%
	Treatment	0.17*	0.02	32%	44%
Sports Materials	Comparison	0.06		28%	32%
	Treatment	0.11	0.06	28%	35%
Concrete	Comparison	0.18		24%	38%
	Treatment	0.18	0.49	26%	40%
Polymers	Comparison	0.11		41%	48%
	Treatment	0.31*	0.000004	37%	57%

* *significant at p<0.05*

Students' Gains in Science Knowledge — Standardized Mean Gain Effect

Combined Tests

- Another measure of gain often applied in educational research is the Standardized Mean Gain Effect size. Effect size (ES) is a name given to statistics that measure the magnitude of a treatment effect in terms of standard deviation units. (Table 16)
- Comparing the Combined *MWM* Treatment student score gains with those of the Comparison Group, there is **preliminary evidence to indicate a respectable significant gain effect size (p<0.00004)** with a standardized mean gain effect of 0.72 (CI= +/-.10) for the *MWM*-taught students. In contrast, the Comparison Group's Standardized Mean Gain Effect was 0.51 (CI= +/-.11).

Table 16: Students' Gains in Science Knowledge – Overall Standardized Mean Gain Effect

Module	*Group*	*Standardized Mean Gain Effect*	*95% Confidence Interval +/-*	*t-test*
Combined Tests	Comparison	0.51	0.11	0.00004
	Treatment	0.72**	0.10	

** significant at p<0.05*

Table 17: Students' Overall Gains in Science Knowledge – by Module Standardized Mean Gain Effect

Module	*Group*	*Standardized Mean Gain Effect*	*95% Confidence Interval +/-*	*t-test*
Composites	Comparison	0.65	0.32	0.045
	Treatment	1.01**	0.22	
Sports Materials	Comparison	0.37	0.25	0.08
	Treatment	0.58	0.24	
Concrete	Comparison	0.94	0.26	0.38
	Treatment	0.80	0.24	
Polymers	Comparison	0.37	0.18	0.000004
	Treatment	1.05**	0.26	

** significant at p<0.05*

Science Knowledge Gains by Demographics

The gains of the Treatment group students varied across demographic groups.

Gender

- For the same time investment, both **boys and girls learning with** *MWM* **significantly out performed their Comparison peers**, as measured by the Percent Value Added, normalized gain and standardized mean gain effect. (Table 18)

Percent Value Added (PVA)

- The Percent Value Added suggests that Girls taught with *MWM* improved their knowledge of science twice as much as Girls taught with regular classroom approaches.
- **Treatment Girls gained an average of twice as much science knowledge** as measured by the pre- to post-test score gains, relative to the Comparison Girls.
- Learning with *MWM*, the Treatment Girls improved their science knowledge an average of 41% pre-to-post-test relative to the Comparison Girls, averaging 21%, which is statistically significant ($p<0.007$).
- The Treatment Boys, too, made stronger knowledge gains with a 42% average test score improvement relative to the Comparison Group Boys with an average 32% test score gains. This gain is a statistically significant difference at the $p<0.005$ level.

Normalized Gain

- As measured by normalized gain, for the same time investment (15 hours), both male and female students taught with *MWM* learned significantly ($p<0.04$) more science than their peers in the Comparison Group.

Standardized Mean Gain Effect

- The standardized mean gain effect for the Treatment Girls was a respectable 0.66 (CI +/-0.21) versus 0.38 (CI +/- 0.18) for the Comparison girls.
- For the males, the standardized mean gain effect is a respectable 0.81 vs. 0.65, Treatment vs. Comparison, respectively.

Table 18: Students' Gains in Science Knowledge – Gain Data by Treatment and Sex

	Percent Value Added	*t-test*	*Normalized Gain*	*t-test*	*Standardized Mean Gain Effect*	*95% Confidence Interval*
Comparison Males	32%		0.21		0.65	0.19
Treatment Males	42%*	p<0.005*	0.27*	0.04	0.81	0.18
Comparison Females	21%		0.14		0.38	0.18
Treatment Females	41%*	p<0.007*	0.22*	0.04	0.66	0.21

** significant at p<0.05*

Race

- ***MWM* significantly improved both White and Non-White students' knowledge gains.**

Percent Value Added (PVA)

- The preliminary evidence suggests **that both White and Non-White Treatment students learned significantly more science knowledge** than Comparison students of the same racial groups. (Table 19)
- **Non-White Treatment students performed better (37% PVA) than White Comparison students (27%) as well as their Non-White counterparts' (25%)**, which is approaching statistical significance at the p<0.06 level

Normalized Gain

- **Looking at the value added by *MWM*, given the same 15 hours of exposure, both Treatment White and Non-White students out performed the Comparison racial groups.**
- Non-Whites in the Treatment Group improved 4% more than those in the Comparison Group. (20% versus 16% normalized gain).
- Whites in the Treatment Group significantly out performed by 8% Comparison White students. (28% vs. 20% normalized gain, p<0.03).

Standardized Mean Gain Effect

- As measured by the standardized mean gain effect, learning via *MWM* had considerable, but not significant, effect on the White Treatment students, who improved their science knowledge by 0.84 vs. 0.51 standardized mean gain effect for the Comparison students).
- For Non-White students, the standardized mean gain effect for each is 0.59 vs. 0.53 Treatment vs. Comparison, respectively, which was not significantly different.

Table 19: Students' Gains in Science Knowledge – Gain Data by Treatment and Race

	Percent Value Added	***t-test***	***Normalized Gain***	***t-test***	***Standardized Mean Gain Effect***	***95% Confidence Interval +/-***
Comparison Whites	27%		0.20		0.51	0.19
Treatment Whites	44%*	p<0.001*	0.28*	0.03	0.84	0.19
Comparison Nonwhites	25%		0.16		0.53	0.19
Treatment Nonwhites	37%	p<0.06	0.20	0.17	0.59	0.19

* *significant at p<0.05*

Grade Level

- Given the same amount of instructional time, **both High School and Middle School students learning with *MWM* out performed their Comparison peers in knowledge gains.**

Percent Value Added (PVA)

- ***MWM* provided significant "value-added" over and above traditionally taught science students.**
- **High School Treatment students made greater science knowledge gains than Middle School Treatment students.**
- The *MWM*-taught students – both Middle School as well as High School – performed significantly better ($p<0.05$ and 0.008, respectively) on pre- to post-test gains than did the Comparison Students. (Table 20)
- As a group, the *MWM*-taught Middle School students demonstrated 12% more knowledge gain pre- to post-test than the Comparison Middle Schoolers: 41% vs. 29% gain, respectively, significant at the $p<0.05$ level.

Normalized Gain

- While both Treatment and Comparison High School students improved their knowledge, Treatment High Schoolers learned significantly (11%, $p<0.007$) more science during the same time period (34% versus 23% normalized gain). (Table 20)
- Middle School students taught science with *MWM* also out performed their Comparison Group peers, (20% to 15% normalized gain) at a rate approaching statistical significance ($p<0.07$).

Standardized Mean Gain Effect

- *MWM*-taught students, whether Middle School or High School, demonstrated a larger Standardized Mean Gain Effect in comparison to the Comparison Students, although not statistically significant. (Table 20)
- The Standardized Mean Gain Effect size for the Middle School Treatment students was 0.69 vs. 0.54 for the Comparison Group.
- The *MWM* High School students' Standardized Mean Gain Effect was 1.04 compared to 0.70 for the Comparison Group.

Table 20: Mean Gain Data by Treatment and Grade

	Percent Value Added	*t-test*	*Normalized Gain*	*t-test*	*Standardized Mean Gain Effect*	*95% Confidence Interval +/-*
Comparison Middle School	29%		0.15		0.54	0.18
Treatment Middle School	41%*	p<0.05	0.20	0.07	0.69	0.20
Comparison High School	31%		0.23		0.70	0.25
Treatment High School	41%*	p<0.008	0.34*	0.007	1.04	0.29

* *significant at p<0.05*

MWM's Impact on Students' Science Attitudes and Skills

Introduction

Due to the two-tiered random selection and assignment of the sample, the results of the Pre-Institute Survey demonstrate that **both the Treatment and Control Groups entered the Institute with similar self-reported attitudes toward science study, careers and similar problem-solving and scientific inquiry skills and behaviors.**

Changes in students' attitudes and behaviors Pre- to Post-Institute were assessed through a battery of self-report questions, administered to students, and calculated into an Overall Score and three subscales scores. A fourth subscale assessed the students' perception of their classroom learning experience. Specifically, the subscales assessed:

Subscale 1: Students' attitudes toward science study and career aspirations;

Subscale 2: Students' ability to use science skills including scientific inquiry to problem solve;

Subscale 3: Students' teamwork skills;

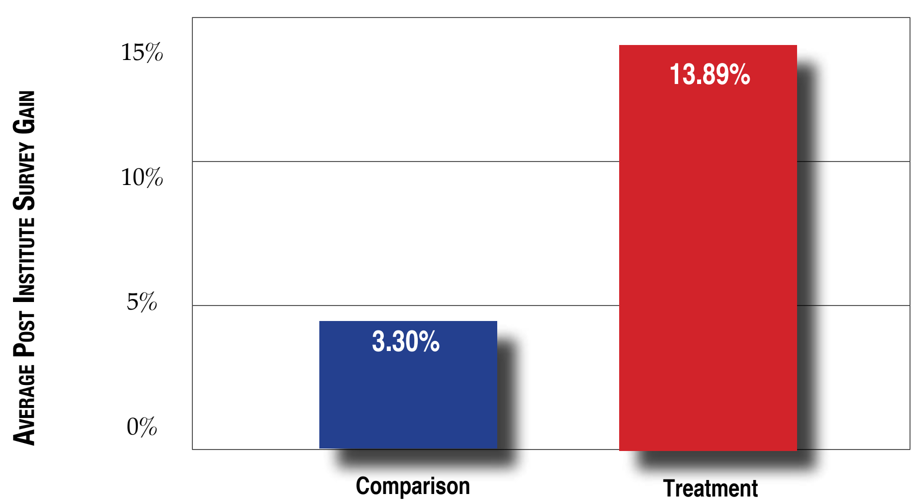
Subscale 4: Students' perceptions of whether they were taught scientific inquiry and problem-solving skills during the Summer Institute.

Students' Self-Reported Attitudinal and Skill Changes

Overall, Treatment students learning with *MWM* reported improving their science-related attitudes and skills more dramatically than did Comparison students over the course of the Institute.

The difference in average positive change (13.89 average vs. 3.30) in the Post-Institute survey approaches a level of statistical significance ($p<0.07$). (Figure 5)

FIGURE 5
STUDENTS' SELF-ASSESSED GAINS IN ATTITUDES AND SCIENCE SKILLS



BASE: PRE-INSTITUTE SURVEY SCORES

1. Treatment students reported larger pre- to post- instructional **improvements on every science-related attitudinal and skill subscale** relative to the Comparison Group students, given the same number of hours of instruction**.** (Table 21)

2. Treatment Group students reported more **strongly improved attitudes towards studying science** (2.45 vs. 0.25). (Table 21)

3. Treatment students reported more improvement in their **ability to use inquiry skills and to problem solve**, than did the Comparison Group (6.58 vs. 1.78). (Table 21)

4. Treatment students also reported their **team work skills improved** much more than the Comparison Group students (0.74 vs. -0.65). (Table 21)

Table 21: Changes in Students' Self-Assessed Science-Related Attitudes and Skills

Scales	***Comparison Gains***	***Treatment Gains***	***t-test p<***
Subscale 1: Students' Self-Assessed **Attitudes Toward Science Study** *(12 items)*	0.25	2.45	0.11
Subscale 2: Students' Self-Assessed Improvement in **Inquiry Skills** and **Problem Solving**. *(19 items)*	1.78	6.58	0.07
Subscale 3: Students' Self-Assessed Improvement in **Team Skills**. *(5 items)*	-0.65	0.74	0.07
Total Scale: Overall Improvement in Attitudes and Skills *(36 items)*	3.30	13.89	0.07

Students' Self-Reported Attitudinal and Skill Changes – Selected Items

Table 22 summarizes the three attitude and skills subscale scores and highlights several survey items on which the Treatment students made much higher gains in improved attitudes and science skills than did the Comparison Group. Treatment students scored significantly higher on all the listed items below, with one exception. Comparison students agreed at a higher rate to the statement: "Laboratory science is boring." Treatment students did not agree with this statement. The difference was statistically significant (p<0.04).

Table 22: Changes in Students' Self-Assessed Science-Related Attitudes and Skills – Selected Items

Attitudes, Skills, Behaviors	*Comparison*	*Treatment*	*t-test p<*
Subscale 1: Attitudes Towards Science *(12 items)*	0.25	2.45	0.12
Science classes are interesting.	-0.28	0.24*	0.02
Laboratory science is boring.	0.30	-0.18*	0.04
I enjoy doing science experiments.	-0.08	0.34*	0.04
Subscale 2: Inquiry Skills and Problem-Solving (*19 items)*	1.78	6..58	0.07
My teacher asks questions to stimulate me to come up with my own answers.	-0.13	0.50*	0.01
Science classes encourage me to discuss my ideas.	-0.08	0.45*	0.05
I develop scientific explanations following rules of logic and evidence.	-0.20	0.29*	0.03
I provide alternate explanations to solve a problem.	-0.18	0.21*	0.06
I apply concepts/ideas I've learned in real-world design problems.	0.08	0.63*	0.03
I design useful things in science class.	0.13	0.55*	0.04
I design a test of the product/project or lab.	0.15	0.61	0.07
Subscale 3: Teamwork Skills *(5 items)*	-0.65	0.74	0.07
Be more inquisitive.	-0.35	0.08*	0.04
Be more self-reliant and take charge of my own learning.	-0.25	0.32	0.06
Total Scale: *(36 items)*	**3.30**	**13.89**	**0.07**

* *significant at p<0.05*

Students' Perceptions of Classroom Teaching Behaviors

Consistent with and reflective of the intent and design of the Summer Institute, Treatment students were far more likely to report being taught to use scientific inquiry and problem–solving skills in their science classes at the Summer Institute than were the Comparison Group Students. (Table 23) The students were asked their perceptions as to which skills were taught in the classroom to:

1) serve as an additional check on the implementation of the classroom intervention, and

2) to provide another measurement to ascertain the impact of the teaching experience on students, in addition to changes in science knowledge scores.

Table 23: Students' Reported Frequency of Inquiry Teaching

Scales	*Comparison*	*Treatment*	*t-test p<*
Subscale 4: Students' Observed Frequency of Being Taught Science Inquiry and Problem-Solving Skills at Summer Institute *(16 items)*	0.08	1.30	0.07

Teachers' Self-Reported Teaching Behaviors and Assessment of Students' Skills Changes

Self-Reported Teaching Behavior

As part of the process to establish that the intervention was successfully implemented, and as part of the data triangulation process, teachers were asked at the end of every module, every week to assess how often they used key teaching behaviors in their classrooms during the week, applying a 5-point Likert scale, with *1= Never and 5 = Extensively*. A series of statements were developed for the Post-Institute Survey, with twelve statements representing the "Desired Treatment Teaching Behavior Scale," and eleven statements representing the "Desired Comparison Teaching Behavior Scale." The average was computed for each scale, to equalize the number of items. Desired Treatment Teaching Behaviors included such acts as the following:

"Required/encouraged students to come up with alternative explanations."

"Required/encouraged students to justify their conclusions."

The "Desired Comparison Teaching Behaviors Scale" included such statements as:

"Stated the conclusions for the students."

"Told students when they were wrong."

Treatment and Comparison Teaching Behavior statements were randomly mixed into a single list. The teachers' responses to the statements were subsequently analyzed by subscale and the results are presented in Table 24 .

The data indicate that the teachers reported predominantly using the desired teaching behaviors of their assigned treatment group. They did so at a statistically significant level. It appears that the Treatment Teachers understood their role more clearly than the Comparison Teachers ($p<0.000001$ vs. 0.0004), although both scores were highly significant. (Table 24)

According to these data, both sets of teachers, therefore, complied with the parameters of the study. These data were also verified against the students' observations of teaching behaviors (Table 23) and the outside researcher's independent observations. (See "Discussion Section" of report)

Table 24: Teachers' Self- Assessed Teaching Behavior

Scales	*Comparison Teachers*	*Treatment Teachers*	*t-test p<*
Teachers' Self- Assessed Teaching Behavior			
Treatment Behaviors *(12 items)*	3.02*	3.90*	0.045
Comparison Behaviors *(11 items)*	4.01*	2.31*	0.00001
t-test (p<)	0.0004	0.000001	

* *significant at p<0.05*

Teachers' Assessments of Changes in Students' Skills

Teachers assessed students' improvement at the end of every module and every week regarding Science Inquiry, Problem Solving, and Team Skills using a 5-point Likert scale with *1= No improvement*, and *5= Big Improvement.*

The Treatment Teachers reported larger improvements in skills for their students on all three subscales than did the Comparison Teachers, although there is not a statistically significant difference. (Table 25)

These data were also verified against the students' self-assessment of changes in their skills and the outside researcher's independent observations. (See "Discussion Section" of report)

Table 25: Subscale Scores Summarizing Teachers' Assessments of Changes in Students' Skills

Subscales	***Comparison Teachers***	***Treatment Teachers***	***t-test p<***
Observed Improvement in Students' Science Inquiry Skills and Problem-Solving *(19 items)*	2.53	3.45	0.11
Observed Improvement in Students' Teamwork Skills *(5 items)*	2.78	3.39	0.18
Observed Students as being More Engaged in Learning *(1 item)*	2.75	3.71	0.10

* *significant at p<0.05*

Discussion

Adequacy of Study Sampling, Intervention, and Instrumentation

In this section, the adequacy of the sampling process, implementation of the intervention and instrumentation is examined and discussed.

Adequacy of Sampling

The final sample of 82 students was equal in size and demographic distribution across the Treatment and Comparison Groups. The sample was generally representative of the broad range of types of middle school and high school students found in Maryland and reflective of the state's demographic proportions and locales. The two-tiered random selection of students and assignment to treatment group appears to have been effective in distributing students equally and proportionally across the two treatment groups, and in minimizing or eliminating several sources of bias. This was demonstrated by the fact that the Pre-test scores and the Pre-Institute Survey scores for both treatment groups were virtually identical. Therefore, whatever changes occurred can be attributable to the intervention that was implemented during the Summer Institute. In the next phase of the evaluation, the sample will be expanded to 1,000 - 3,000 Treatment students.

Adequacy of the Intervention

The intervention was successfully implemented as evidenced by the fact that the data from the students, teachers and independent researcher all confirm the scope and directionality of the implementation in the Comparison and Treatment classrooms. (Table 26)

Table 26: Comparison of Students' and Teachers' Perceptions of the Intervention

Scales	***Comparison Gains***	***Treatment Gains***	***t-test p<***
Students' Observed Frequency of Being Taught Science Inquiry and Problem-Solving Skills at Summer Institute *(16 items)*	0.08	1.30	0.07
Teachers' Self- Assessed Teaching Behavior: *(23 items)*			
Inquiry-Based Teaching Behaviors *(12 items)*	3.02	3.90*	0.045
Traditional Teaching Behaviors *(11 items)*	4.01	2.31*	0.00001

* *significant at p<0.05*

Adequacy of the Instrumentation

As evidenced by the results and the relatively high number of significant findings, the instruments were sufficiently well-constructed, comprehensive and reliable to 1) identify and capture major changes in students' science knowledge, attitudes and skills, 2) to produce reliable tests, and 3) to triangulate the data to enhance the analysis' reliability and validity.

Addressing the Four Research Questions

A series of significant research questions were developed at the proposal stage to direct the empirical investigation. These questions formed the heart of the research plan and were developed to guide the research work. The research questions reflected the objectives of the evaluation and provided a practical structure and a framework for stakeholders. The following discusses the results of the study relating them to address the research questions.

Discussion of Results Addressing the Four Research Questions

The following discusses the study results addressing the four research questions:

1. ***Compared to students in "regular science classes" with teacher-led, textbook instruction (Comparison Group), what is the impact of MWM Modules (Treatment Group) on improving students' science knowledge?***

   The initial evaluation of the effectiveness of *Materials World Modules* compared with traditional teacher- and textbook-centered approaches, indicates there is respectable *preliminary* evidence to suggest the following:

   - **The *MWM*-taught students improved their science knowledge significantly more than those taught the conventional way.**
   - ***MWM* students made significantly larger science knowledge gains** than those in the Comparison Group **for the same time investment**. Starting the institute with the same average pre-test scores, Treatment students improved their pre- to post-test science knowledge.
   - *MWM* provided significant "value-added" over and above traditionally taught science students, as measured by all three statistical measures—Percent Value Added, Normalized Gain Effect, and Standardized Mean Gain Effect.
   - This represents a statistically significant ($p<0.0001$) difference, suggesting a strong probability that the treatment (*MWM*) was more effective than traditional classroom approaches.
   - *MWM* Modules led to improved results for **both boys and girls** taught with the modules when compared with the Comparison Group. **Treatment Girls' gained an average of twice as much science knowledge** as measured by the pre- to post-test score gains, compared to the Comparison Group Girls.
   - In addition, *MWM* significantly improved **both White and Non-White students'** knowledge gains.
   - ***MWM* helped Non-White students out-perform both White as well as Non-White students** taught the traditional way.
   - Given the same instructional time, **both High School and Middle School students** learning with *MWM* out performed their Comparison peers.

2. ***Compared to students' in regular science classes, what is the impact of the MWM Modules on improving students' use of the scientific method to problem solve and design projects?***

- **Both the students and the teachers who taught them concur that *MWM* improved Treatment students' ability to use scientific inquiry skills in the classroom during the Summer Institute.**
- **Both students and their teachers reported the *MWM*-taught students enhanced their ability to problem solve at a higher rate than the Comparison Group.**
- ***MWM*-taught students report the change in their ability to use inquiry and problem-solving skills to be approximately 3 times higher than the students taught conventionally.**
- Teachers also observed a change in the students, with the teachers of the *MWM* students rating them as more highly improved in using inquiry and problem-solving skills than did the Comparison Teachers.
- However, the teachers did not observe the students' changes to be as dramatic as the students' self-ratings reflect.
- **Comparing students' ability to apply scientific inquiry and problem-solving skills to solve problems on essay tests (rubric questions) and providing adequate scientific explanations, *MWM* students performed at a significantly higher level than did the Comparison Group.**
- The independent researcher also observed that the Treatment students generally used a more systematic approach to problem-solving, and the students reported that as a result of their *MWM* instruction they now feel they can learn positive lessons from science lab errors made.

3. ***Compared to students in regular science classes, what is the impact of the*** MWM ***Modules on improving students' attitudes toward the study of science or careers in the science field?***

- **Treatment and Comparison students entered the Summer Institute with similar attitudes toward science and similar science skills.** A Pre-Institute battery of 53 questions encompassing career aspirations, inquiry skills, problem solving, teamwork, and learning experiences produced virtually the same results for Treatment and Comparison students on the Pre-Institute survey.
- **Overall, Treatment students, teachers and the independent researcher concur that the *MWM* students improved their science-related attitudes and skills more dramatically than did Comparison students** over the course of the Institute.
- **Treatment students and their teachers reported larger** pre- to post-module **improvements on every science-related attitudinal and skill subscale** relative to the Comparison Group students, **given the same number of hours of instruction**.
- **Treatment students and their teachers reported more strongly improved attitudes toward interest in studying science than did the Comparison Group.**
- Treatment students and their teachers also reported the ***MWM* students' teamwork skills improved much more** than did the Comparison Group students.

4. ***Compared to teachers in regular science classes, what is the impact of the* MWM *Modules on improving science teachers' abilities to train students to apply scientific research methods to problem solve*?**

- **Students reported, and teachers concurred, that due to changes in teaching and the MWM materials the *MWM* students improved their ability to apply scientific inquiry methods to problem solve more successfully** than did the Comparison group students, although not at a statistically significant level. The researcher's classroom observations of teachers and students supported this finding.
- In addition, ***MWM* students performed significantly better than their Comparison counterparts on long and short essay questions on the module tests in which they had to use inquiry skills to solve a problem using scientific thinking.**
- However, there were a **number of items on the students' and teachers' scales that indicated there were some areas of instruction that require further attention by the teachers to help students learn to become fully proficient in scientific inquiry and problem-solving.** The independent researcher's classroom observations concurred with these quantitative results.

Recommendations: Next Steps

- A full-scale evaluation of *MWM's* effectiveness should be conducted in regular classroom settings during the school year with a minimum sample of 1000 Treatment Students and 1000 Control Students to be able to project *MWM*'s effectiveness in other educational settings with statistical confidence. The full scale study should be implemented as soon after the *MWM* program is implemented in schools as possible, between 2007 and 2009.
- Based on the rich and multi-layered evaluation data gathered from the 2006 Summer Institute, a full-ranged analysis of the teachers' and students' assessments of their experience with *MWM* should be conducted, with follow-up reports and briefings conducted with *MWM* stakeholders, to form a strong analytical basis for designing and implementing the next stage of the *MWM* roll-out.
- Drawing on the evaluation data and other sources, the teacher-training component should be strengthened and designed to train teachers to better instruct and coach students to become more proficient in every inquiry and problem-solving skill.
- Each *MWM* module should more clearly define and present basic learning objectives to be accomplished in each module to help teachers gain the most from the materials.
- In addition to establishing basic learning objectives, the *MWM* modules should also encourage the teachers to develop their own objectives adapted to their classroom.
- *MWM* should train teachers to know the difference between 1) teacher-directed instruction, 2) inquiry-based instruction, and 3) a "laissez-faire" approach to science teaching and learning which is sometimes mistaken as "inquiry based" instruction.
- *MWM* should train teachers to use inquiry-based instruction correctly to coach their students to know, understand, apply and be able to discuss the underlying science concepts in labs and design projects.
- Teachers should train students to use the *MWM* approach to not only "report lab data" but also to always relate it to the underlying science concepts.
- Teachers should be trained how to effectively "coach" and subtly "lead" students to apply their skills in the laboratory and project design components to maximize the learning and inquiry process.
- Teachers should be trained to consistently challenge their students to answer questions, to challenge each others' scientific thinking and assumptions and to be able to scientifically defend their laboratory and design process and product.
- A full scale review of the literature should be conducted prior to the Formal Evaluation to identify and confirm the variables that contribute to student science achievement and teacher effective teaching of students to implement inquiry- and design-based science.
- The *MWM* project should develop reliable standardized assessments for each of the available modules. This is a multi-step, iterative process that will take two to three years to develop. The standardized tests are critical to the evaluation and increasing the value of *MWM* to school districts and teachers.
- The standardized assessments should be included with the materials as part of each *MWM* module.
- *MWM* should revise the module test items, based on the psychometric data.
- *MWM* should develop new test items for the four modules used during the Summer Institute and the other five modules, for field-testing and psychometric analysis.

Appendix A

Overall *MWM* Evaluation Plan – Three Phases

Action Research & Associates, Inc. is conducting a scientifically rigorous evaluation to serve as the foundation for collecting the data required to evaluate the effectiveness of *Materials World Modules.* Building on its experience conducting major educational research and evaluation projects, Action Research designed an "**Overall *MWM* Research Plan**" to guide the long term evaluation of *MWM* (2006-2009) of implementation of *MWM*. The "**Overall *MWM* Research Plan**" is based on a single scientific research model (see Figure 3) to guide gathering and analyses of data. By applying the same model, research design, methodology, instruments and analyses in three studies over the three years of *MWM's* roll-out in Maryland or other states, it will be possible to scientifically explore, replicate and then generalize regarding *MWM*'s effects on students.

The evaluation model employs a quasi-experimental, mixed-method design, gathering both quantitative and qualitative data. The data sources are "triangulated;" that is, to answer the research questions, multiple data sources are assessed and compared to verify the reliability of the results. For instance, how *MWM* affects students' ability to problem solve will be assessed by student self-assessments, teacher and volunteer assessments of the students, and independent researcher classroom observations.

Thus, the data from the replicated studies will provide a scientific, reliable and valid basis to guide DoD in making a well-reasoned and supported decision regarding whether or not to implement *Materials World Modules* in other locations in Maryland and throughout the Nation.

Overall *MWM* Evaluation Plan

The evaluation research model and the target sample size proposed for the *MWM* study is comparable to a previous study conducted for the U.S. Department of Education by the Principal Investigator[1]. The study was cited by the USDE officials [2], as one that met the federal government's scientifically-based research (SBR) standard. A comparable research model, sample size and framework is being applied in Action Research's "Overall *MWM* Evaluation Plan" to evaluate the *Materials World Modules'* impact on students' science knowledge, behavior, attitudes and aspirations.

To produce a study with reliable results for decision-makers, it will be necessary to attain a minimal sample-size of 987 Treatment students for Maryland and 3,000 Treatment students for the Nation. The sample-size is determined based on a statistical calculation of the sample size required in a quasi-experimental study to be able to conduct power statistical analyses of the data. These data will be necessary to develop reliable analyses and evidence that DoD can use to base their decisions regarding whether or not to proceed to a national scale-up of *MWM*. The following provides a brief description of the proposed Three Phases of the *MWM* Evaluation:

---

[1] *"Data Collection of Federal Performance Indicators for PBS Ready to Learn: Year 4 Summary Report,"* Horowitz, J: K. Juffer; L. Davis; J. Stout; J. Bojorquez; K. Dailey; E. Holms. WestEd, Los Alamitos, CA, August 2004.

"Data Collection of Federal Performance Indicators for PBS Ready to Learn: Year 4 Summary Report," Horowitz, J: K. Juffer; L. Davis; J. Stout; J. Bojorquez; K. Dailey; E. Holms. WestEd, Los Alamitos, CA, August 2004.

[2] *"Current Methodologies Supported by the Federal Government,"* Co-panelists: Juffer, K.A., (WestEd); A. Dixon (U.S. Dept. of Homeland Security), M. Silverberg (U.S. Dept. of Education). T. Knicker, (U.S. Dept. of State), G. Della-Piana (National Science Foundation). Evaluators' Institute and Washington Evaluators Association: Washington, DC, July 2004.

Preliminary Evaluation of *MWM*

- **Phase I Evaluation (July 9-August 5, 2006) *MWM* Summer Science Institute**

  Phase I was conducted by Action Research & Associates, Inc. to meet DoD's initial time line and requirement to provide preliminary data on *MWM*'s effectiveness by Fall 2006. Due to a compressed time line in Phase I, it was necessary to evaluate the *MWM* modules' effectiveness outside the normal school year. To do so, it was decided to develop a controlled learning environment– a four-week 2006 *MWM* Summer Institute – in which to conduct a quasi-experimental evaluation study. Phase I also was designed to be an important step to develop reliable instruments and refine them to meet the study's objectives. It was necessary to initially develop and then test and enhance the reliability and validity of the *MWM* knowledge tests and survey instruments. Action Research scientifically refined them, based on statistics gathered from Phase I to prepare for Phases II – IV. These three phases would form the "Formal Evaluation of *MWM*."

Formal Evaluation of *MWM*

- **Phase II Evaluation (2007-2008) Mid-Sized School Districts**

  This is the first of three planned phases of the Formal Evaluation of *MWM*. In contrast to the intensive four-week immersion program – the *MWM* Summer Science Institute-- in Phase I, the Formal Evaluation of *MWM* will occur in a "naturalistic" setting. For example, in 2007 *MWM* will be tested in a mid-sized Maryland school district during the regular school year in regular science classrooms adjacent to a DoD research laboratory facility, for example, Harford County School District. A new component of this study will be the DoD Volunteers (engineers and researchers from the research labs) and how they affect *MWM*'s impact on students and teachers. Action Research will analyze the item statistics generated from the Phase I study test and surveys. Action Research will then revise the surveys and tests and these instruments will be deployed in Phases II, III, and IV. Phase II in the Mid-Sized District is expected to yield data on approximately 500 students (250 Treatment, 250 Comparison), 20 Teachers (10 Treatment, 10 Comparison), and five DoD volunteers. Depending on availability of funds, other districts like Garrett County (MD) School District or other states may also be added to capture the data on another 480 students and 16 teachers (half Treatment, half Comparison).

- **Phase III Evaluation (2008-2009) - Large School District**

  It is expected that the *MWM* State Center for STEM Education will recruit at least one or more large school districts in Maryland or other states. Assuming the State Center will train teachers in *MWM* from these counties, the same quasi-experimental research design will be applied. Other large school districts may be added to this phase.

Cumulative Study

Cumulative Study: At the end of Phase III**,** Action Research proposes conducting a **"cumulative study"** aggregating and analyzing the evaluation data from Phases II and III, in order to achieve the minimum sample size of 1,000 Treatment students for Maryland and/or 3,000 Treatment students for the Nation. The study will explore the effects of *MWM* on students, teachers and volunteers. In addition to the traditional parametric tests, and other analyses, Action Research will conduct power statistical analyses of the effects of *MWM* on science students using factor analysis and Hierarchical Linear Modeling as well as other statistical techniques, as appropriate. Action Research proposes producing a **"Final Report of the Formal Evaluation of *MWM*"** which will be based on rigorous scientific research and valid and reliable data that has been replicated over three years in three or more school districts in Maryland or other states.

ACTION RESEARCH

Action Research & Associates, Inc.

Dr. K.A. Juffer, Executive Director of Action Research & Associates, Inc. (established in 2000), designed and conducted the *Materials World Modules* evaluations. Dr. Juffer is well-qualified to conduct these evaluations since she has been a professional researcher for more than 25 years, conducting high caliber research and educational program evaluations for such clients as the National Science Foundation, the National Institutes of Health, U.S. Department of Health and Human Services, U.S. Department of Education, the Public Broadcasting Service (PBS), U.S. Information Agency, the Fulbright Program, the Voice of America radio and television programming, Arbitron, Inc, CBS network television and radio, ABC network television, Clear Channel radio and other media companies; Prince George's (MD) County Public Schools, the University System of Maryland's Chancellor's Office and other agencies, corporations, universities and school districts.

In addition, Dr. Juffer is an experienced developer of tests and assessments. She conducted psychometric research for the Iowa Testing Program, a subsidiary of American College Testing, developing reliable and validated assessment instruments and achievement tests. She also developed a highly reliable (.95), nationally-recognized psychological test, the *Culture Shock Adaptation Inventory* (CSAI ©1983), with four unique subscales that contribute to an overall test score measuring the degree of cross-cultural adaptation an individual experiences, which was a break-through in its field.

With masters and doctorate degrees in Education, Dr. Juffer has worked in the education field for more than 25 years, as a classroom teacher, district administrator, curriculum coordinator, university professor, evaluator, contractor and federal government official. She has conducted educational research and communications research for more than 25 years for local, state, national and international clients.

Acknowledgements

Special acknowledgement and appreciation is extended to Dr. Barbara Pellegrini, researcher at Northwestern University working on the development of *Materials World Modules* for her professional and highly collegial guidance and generous sharing of materials, and professional insights. I particularly want to gratefully acknowledge the significant statistical work and professional and personal support Dr. Steve Greenfield has contributed. His professional guidance, insights, friendship, and tireless devotion to cleaning databases and crunching data made the many hours of working on this project much more interesting and enjoyable, contributing to the final product.

I also want to acknowledge the efforts of John Yochelson of Building Engineering and Science Talent (BEST), whose support, guidance and insights significantly helped polish this report and the Executive Summary. I value the professional friendship and camaraderie we have developed. A special thanks is extended to Dr. Stephen Herman, former President of Garrett College, for his valuable support and insights in producing this report. I also want to acknowledge the professional guidance and support of Deborah Yourick, Deputy Director of the Walter Reed Army Institute of Research, for her timely and valuable support to guide the Institutional Review Board approval process. A spotlight of gratitude must be shown on Brenda Sullivan of BEST for her creative abilities in preparing and formatting this report for publication.