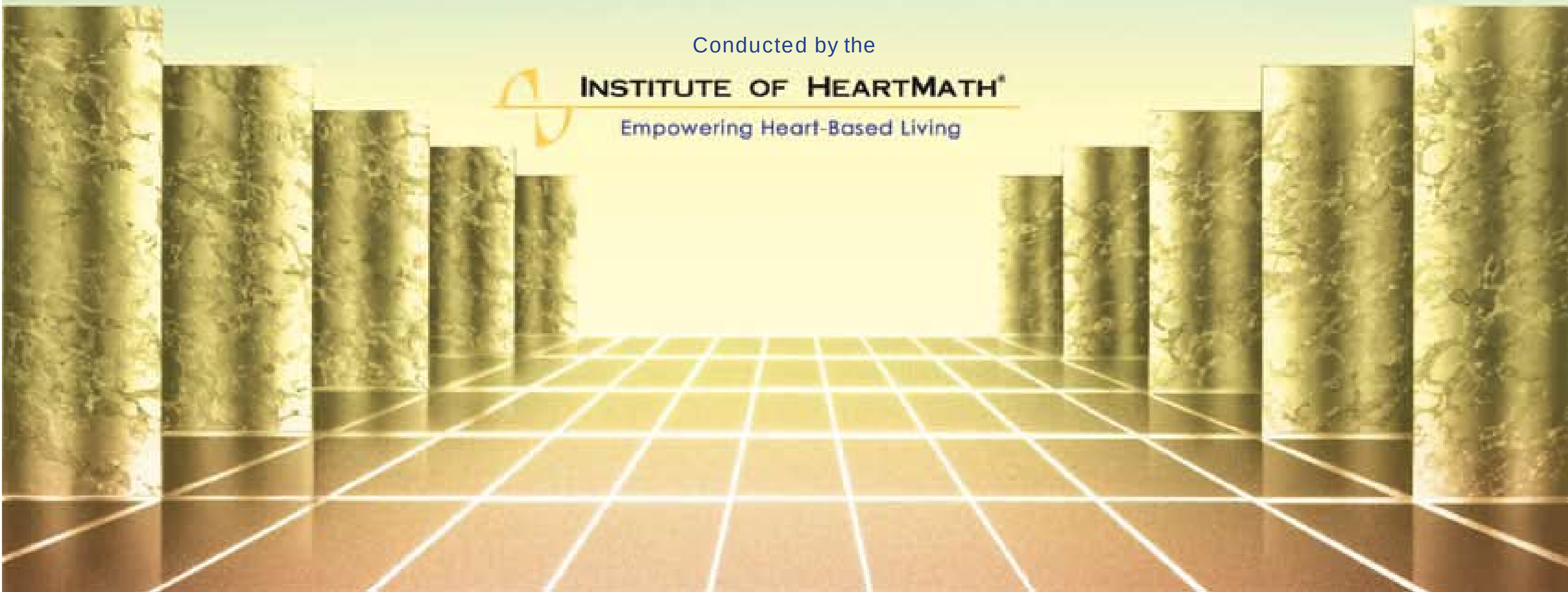
Reducing Test Anxiety and Improving Test Performance in America's Schools

Results from the TestEdge® National Demonstration Study

Raymond Trevor Bradley, Ph.D.
Rollin McCraty, Ph.D.
Mike Atkinson
Lourdes Arguelles, Ph.D.
Robert A. Rees, Ph.D.
Dana Tomasino

Conducted by the

Institute of HeartMath®
Empowering Heart-Based Living



Published in the United States of America by:
Institute of HeartMath
14700 West Park Avenue, Boulder Creek, California 95006
831-338-8500
info@heartmath.org
www.heartmath.org

Cover design and layout by Sandy Royall

Citation information:

Bradley, R. T., McCraty, R., Atkinson, M., Arguelles. L., Rees, R. A., & Tomasino, D. (2007). *Reducing Test Anxiety and Improving Test Performance in America's Schools: Results from the TestEdge National Demonstration Study*. Boulder Creek, CA: HeartMath Research Center, Institute of HeartMath, Report No. 07-04-01.

Address for correspondence:
Rollin McCraty, Ph.D., Director of Research
HeartMath Research Center
Institute of HeartMath
14700 West Park Avenue
Boulder Creek, California, 95006

Ph. (831) 338-8500, Fax: (831) 338-1182, Email: rollin@heartmath.org

Note:

Freeze-Frame, HeartMath, Heart Lock-In, and The Resilient Educator are registered trademarks of the Institute of HeartMath. TestEdge is a registered trademark of HeartMath LLC. Freeze-Framer and emWave are registered trademarks of Quantum Intech, Inc. Attitude Breathing is a registered trademark of Doc Childre.

Recent advances in neuroscience are highlighting connections between emotion, social functioning, and decision making that have the potential to revolutionize our understanding of the role of affect in education. In particular, the neurobiological evidence suggests that the aspects of cognition that we recruit most heavily in schools, namely learning, attention, memory, decision making, and social functioning, are both profoundly affected by and subsumed within the processes of emotion … .

—Mary Helen Immordino-Yang & Antonio Damasio

"We Feel, Therefore We Learn: The Relevance of Affective and Social Neuroscience to Education." *Mind, Brain, and Education*, 2007, 1(1): 3.

Endorsements

"I was thrilled to read HeartMath's comprehensive report on the results of the TestEdge National Demonstration Study. The study is superb, utilizing an experimental design in which data were gathered from a rich combination of questionnaires, interviews, observations, student drawings, and physiological recordings. All of the steps of the study are adequately controlled and the rigorous multivariate statistical procedures are highly effective. The study yielded an impressive body of cross-corroborating evidence documenting the effectiveness of the TestEdge program in reducing student test anxiety and improving test performance. The study is an exemplar of how social science experiments in open field research settings ought to be done.

From my vantage point, obtaining objective physiological recordings of heart rhythm activity from students not only is a significant achievement in its own right, but also adds a new scale of validation to the study's major findings and enhances the value of the study well beyond its immediate application. Of particular import is the physiological evidence indicating that students in the program had established a new set point of emotional stability—a requisite for sustained behavioral change."

—KARL H. PRIBRAM, M.D., PH.D. (HON. MULTI.)
Neuropsychologist; Author – *Brain and Perception* and *Languages of the Brain*;
Professor Emeritus, Stanford University; Distinguished Research Professor,
Georgetown University

"The detailed information presented in this report describes the Institute of HeartMath's comprehensive study on reducing stress and test anxiety and improving students' academic performance and emotional well-being ... . The findings of this study provide impressive evidence that students who were trained to use the HeartMath TestEdge program showed significant reductions in test anxiety and corresponding improvement in their scores on standard measures of academic performance. I was impressed with the careful statistical analyses of the data, and was especially pleased to note the large alpha coefficients for the Total, Worry and Emotionality scores of the Test Anxiety Inventory, and that the scores on these measures were substantially reduced by the TestEdge intervention. While it is unfortunate that the demographic characteristics of the students in the experimental and control schools were somewhat different, the statistical procedures for minimizing the influence of these differences were very effective. Overall, the studies conducted by the Institute of HeartMath provide impressive evidence that the TestEdge program is an effective intervention for reducing test anxiety and facilitating academic performance."

—CHARLES D. SPIELBERGER, PH.D., ABPP
Psychologist; Author – *Understanding Stress and Anxiety, State-Trait Anxiety Inventory,*
and *Test Anxiety Inventory*; Distinguished Research Professor of Psychology and Director,
Center for Research in Behavioral Medicine and Health Psychology,
University of South Florida

"As an author primarily concerned with the development of intelligence in children, whose eight books on that general theme have been published and translated into many languages, … I predict that within a decade nothing less than a true revolution in child development, rearing, and education will be brought about by HeartMath's TestEdge research.

Of all the HeartMath reports, this one holds the greatest promise. It is a goldmine of information, research, and insight ... . That the biological basis for learning lies in the emotional structures of brain and heart, not in intellectual schema and enforced modifications of behavior, is literally the liberation of childhood and the society as a whole. All I can do is applaud, with a grateful heart."

—JOSEPH CHILTON PEARCE
Specialist in early child development; Author – *Magical Child, Evolution's End* and *The Biology of Transcendence*

"The report on the study of the Test Edge program is impressive and convincing. The finding that the TestEdge program led to improvements in student emotional awareness, emotional management, and classroom interactions makes this an effective method for handling challenging and stressful situations such as the current emphasis on testing in U.S. schools. It is impressive that the program not only reduced test anxiety but also improved test performance. It is even more impressive that this technique can transfer to other situations, providing a tool for facilitating the development of positive emotions, cooperative relations, and more effectiveness as well as personal satisfaction."

—RIANE EISLER
Cultural historian; Author – *Tomorrow's Children: Partnership Education for the 21st Century* and *The Chalice and the Blade*

"This is a very well designed and carefully executed study using state of the art assessment techniques confirming that emotional stress impairs academic performance. It also demonstrates the ability of the TestEdge intervention to significantly improve test scores by reducing test anxiety based on personal report as well as objective heart rate variability measurements. School failure was described in a recent interview and book as the leading cause of premature death in the world and is particularly high for minority groups in the U.S. As noted, "The incidence of all types of heart disease increases as education decreases and this stunning relationship is not due to increased poverty or poorer access to medical care." This underscores the need for new innovative approaches such as those described in this study that can improve not only health but also the quality of life by reducing stress."

—PAUL ROSCH, M.D., F.A.C.P.
President, The American Institute of Stress; Clinical Professor of Medicine and Psychiatry, New York Medical College

"The Institute of HeartMath's TestEdge National Demonstration Study accomplishes its primary purpose of testing the efficacy of a program designed specifically to improve the emotional well-being and academic performance of public school students. Utilizing a more robust, diverse sample population, the study validates previous research on the positive impact of the TestEdge program on test scores, passing rates, and psychosocial functioning. The results clearly show that student affect has a largely unrecognized and underappreciated impact on the variance in student performance.

Importantly, the study finds that HeartMath tools produce long-lasting effects on students' psychosocial functioning, well past the intervention period. This finding is especially encouraging when viewed across the range of programs in which students may learn new skills to encourage emotional and physiological coherence. One of the study's surprises is that girls' responsivity to test stress and a test reduction strategy is different from that of boys. This strongly advances the debate for strength-based approaches to teaching/learning in the classroom. Overall, the findings of the Test Edge study motivate educators to pursue the highest level of thoughtful, evidence-based instructional methodologies in our public schools."

—Jorge Calzadilla
Director, Youth Learning Institute, Clemson University

"Stress and anxiety are ubiquitous in modern society, perhaps nowhere more so than among America's youth. Children and adolescents are faced, not only with developmental challenges inherent to their age, but also the challenges of living in a fast-paced world with instant electronic access to world wide news and opinions through social networks such as MySpace and FaceBook. In addition, students face social and academic tests in person at school and in their neighborhood. Families are less stable and more dynamic than a generation ago, and local supportive networks through clubs, churches and civic groups are declining, while formal mental health services are inadequate in most communities. Given the increased challenge and decreased availability of stable social support, it is no surprise that students experience high levels of stress and anxiety.

This combination makes emotional self-management tools a high priority for educators and clinicians. This report contains hopeful and inspiring data about the effectiveness of one such program, TestEdge. Both quantitative and qualitative data in large numbers of students suggest impressive improvements in stress, anxiety and test scores. These results suggest that this program should be implemented widely to enhance the health and achievements of our next generation."

—Kathi J. Kemper, M.D., M.P.H.
Caryl J. Guth Chair for Holistic and Integrative Medicine, Professor, Pediatrics and Public Health Sciences, Wake Forest University School of Medicine

Executive Summary

This Executive Summary provides an overview of the purpose, research methods, and major findings of the TestEdge National Demonstration Study conducted by researchers at the Institute of HeartMath in collaboration with faculty and graduate students at Claremont Graduate University. The Summary has been designed with section headings to assist the reader in locating the appropriate chapters or sections within the main body of the report.

Study's Purpose

The study's primary purpose was to investigate the efficacy of the TestEdge program in reducing stress and test anxiety and improving emotional well-being, quality of relationships, and academic performance in public school students. This involved determining the magnitude, correlates, and consequences of stress and test anxiety in a large sample of students and investigating the degree to which an intervention with TestEdge had a positive effect on students in an experimental group when compared to those in a control group. A second programmatic purpose was to characterize the implementation of the program in relation to its receptivity, coordination, and administration in a wide variety of school systems with diverse cultural, administrative, and situational characteristics.

TestEdge Program

For the purposes of this study, both teachers and students received instruction in HeartMath tools and techniques through the Resilient Educator and TestEdge programs, respectively. Used as the intervention in this study, these programs were developed by the Institute of HeartMath to help students and teachers reduce stress and test anxiety, improve test and academic performance, and enhance emotional and relational competence. The programs are based on 15 years of scientific research on the psychophysiology of learning and performance, emotional dynamics, and heart–brain communication. The programs teach a set of easy-to-use positive emotion refocusing and restructuring techniques that enable teachers and students to self-regulate stress, test anxiety, and other emotional impediments to learning and performance.

The basis of the effectiveness of the techniques is that they enable the individual to self-activate a specific, scientifically measurable psychophysiological state of optimal function, termed *psychophysiological coherence*. Research has shown that psychophysiological coherence is characterized by increased synchronization in nervous system activity, increased emotional stability, and improved cognitive and task performance.

The TestEdge program has been successfully implemented in schools throughout the U.S. and in some foreign countries; pilot studies have shown associated improvements in student standardized test scores, passing rates, and psychosocial functioning. This study marks the first time that the efficacy of the TestEdge program has been evaluated in a large-scale implementation.

Study's Hypotheses

The study tested two major hypotheses. The first is that enhanced competence in emotional management through learning and practicing the TestEdge tools would result in significant improvements in student emotional self-regulation and psychophysiological coherence. These changes then would produce a marked reduction in test anxiety, which, in turn, would generate a corresponding improvement in academic and test performance. Secondly, as a result of the improvement in student emotion regulation skills, it was also expected that there would be associated improvements in stress management, emotional stability, relationships, and overall student well-being, as well as in classroom climate, organization, and function.

Research Design and Methods

To investigate the veracity of these hypotheses, two studies were conducted, each with different research objectives and designs. The first, the primary study, focused on an in-depth investigation of students at the tenth grade level. It was designed as a quasi-experimental, longitudinal field study involving pre- and post-intervention panels of measurement within a multi-methods framework.

For the primary study, extensive quantitative and qualitative data were gathered using survey questionnaires, interviews, structured observation, and an assessment of student drawings, along with student test scores from two California standardized tests—the California High School Exit Examination (CAHSEE) and the California Standards Test (CST). In addition, an electrophysiological sub-study was conducted on a randomly stratified sample of students from both schools. Utilizing measures of heart rate variability (beat-to-beat change in heart rate), this controlled laboratory experiment investigated the degree to which students had learned the techniques taught in the TestEdge program by providing an objective measurement of their ability to shift into the psychophysiological coherence state prior to taking a stressful test.

The secondary study consisted of a series of qualitative investigations to evaluate the accessibility, receptivity, coordination, and administration of the program across elementary, middle, and high schools and in school systems with diverse ethnocultural, socioeconomic, administrative, and situational characteristics. We employed a case study approach to evaluate the implementation of the TestEdge program in nine schools in eight different states (California, Delaware, Florida, Ohio, Maryland, Texas, Wisconsin, and Pennsylvania). Age-appropriate versions of the TestEdge program were delivered to selected classrooms, covering grades 3 through 8 and grade 10. Observational and interview data were gathered to provide information on best practices and potential difficulties when implementing interventions such as TestEdge in widely diverse school settings.

Research Sites and Participants

The primary study involved the entire tenth grade populations of two large high schools in Northern California. (To preserve confidentiality, the identities of the schools are not disclosed in this report.) One high school was randomly selected as the intervention school, while the other served as the control school. While the selected schools were matched as closely as possible on academic performance and key sociodemographic factors, there were some differences in the measurement of academic performance and ethnicity at baseline. Altogether, a total of 980 students participated in the primary study, of which 636 (53% male, 47% female) were in the experimental group and 344 (40% male, 60% female) were in the control group.

Intervention

The Resilient Educator program was delivered to teacher participants in a one-day workshop to provide them with a working familiarity with the HeartMath tools and techniques prior to their beginning classroom instruction of the TestEdge program. Teachers then delivered the TestEdge program to students during the Spring 2005 semester. In the program, students learned and practiced specific emotional management techniques to aid them in more effectively handling stress and challenges, both at school and in their personal lives. They were also taught how to apply these techniques to enhance various aspects of the learning process, including test preparation and test-taking. Both the student and teacher programs included use of the Freeze-Framer (now emWave PC) technology, a heart rhythm coherence feedback system designed to facilitate acquisition and internalization of the emotional management skills taught in the program.

Major Findings: Primary Study

The primary study produced a number of important findings, which are summarized below.

Pre-Intervention Findings

Across the whole sample at baseline, before the TestEdge intervention, the primary study found that:

- Sixty-one percent of all students reported being affected by test anxiety, with 26% experiencing high levels of test anxiety often or most of the time.
- Twice as many females experienced high levels of test anxiety, as compared to males.
- There is a strong negative relationship between test anxiety and test performance; students with high levels of test anxiety scored, on average, 15 points lower on standardized tests in both Mathematics and English-Language Arts (ELA) than students with low test anxiety (Figure 1).
- Five common factors were found in regression analysis models to explain student test performance on the CAHSEE and CST: Test Anxiety-Worry, Test Anxiety-Emotionality, Feelings about School, Life Preparedness, and Educational Plans. For both tests, the regression models explained about 20–24% of the variance in student test performance; test anxiety accounted for about half of the explained variance.

- Multiple regression analysis found that measures of Affective Mood explained approximately twice the variance in student test performance on both the CST-ELA and CAHSEE-ELA as items from the Test Anxiety scale (23% versus ~13%, respectively). Positive feelings and prosocial behaviors have a positive effect on test performance, while strongly negative feelings and antisocial behaviors have a negative impact (Hartnett-Edwards, 2006).

Taken as a whole, these findings are sobering and justify the concern that test anxiety may significantly jeopardize assessment validity and therefore may constitute a major source of test bias.

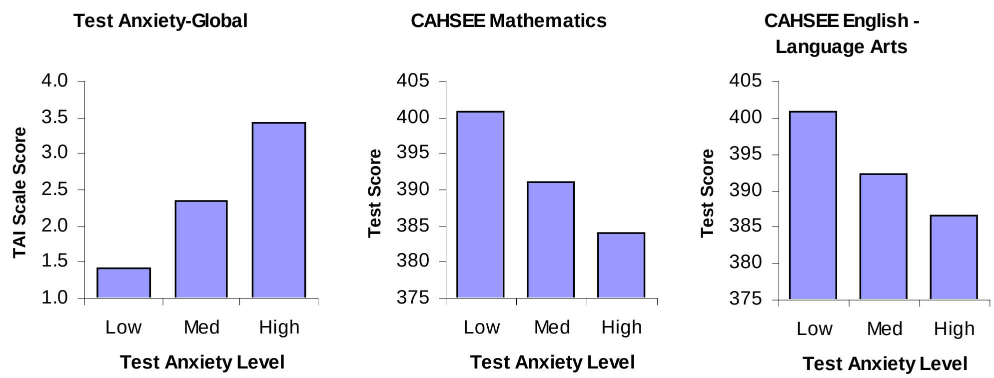


Figure 1. High School Exit Examination Scores by Baseline Test Anxiety Level
Baseline test anxiety, measured by the Test Anxiety Inventory (TAI)-Global Scale score, and midterm California High School Exit Examination (CAHSEE) scores in English-Language Arts and Mathematics have been classified into three approximately equal-sized groupings (tertiles) of students with low, medium, and high test anxiety scores. A strong, statistically significant ($p < 0.001$) negative relationship is clearly apparent between mean level of test anxiety and mean performance on the standardized tests: as test anxiety increases, test performance decreases.

Post-Intervention Findings

After the TestEdge program had been delivered to the students in the experimental school, we found strong, consistent evidence of a positive effect of the intervention on these students when compared to those in the control school:

- There was a significant reduction in the mean level of test anxiety. Of those students at the intervention school who had reported being affected by test anxiety at the beginning of the study, 75% had reduced levels of test anxiety by the end of the study.

- This reduction in mean test anxiety was also evident for more than three-quarters of all classrooms, and it was observed throughout the academic ability spectrum—from high test-performing classes to low.
- The reduction in test anxiety was associated with the following improvements in socioemotional measures (Figure 2):
  - A reduction in Negative Affect (feelings of stress, anger, disappointment, sadness, depression, and loneliness);
  - A reduction in Emotional Discord, reflecting increased emotional awareness and improved emotional management;
  - A reduction in Interactional Difficulty, reflecting increased empathy and improved relations with others;
  - An increase in Positive Class Experience, reflecting perception of increased enjoyment and learning in class, positive feelings toward classmates, and teacher care.
- In four matched-group comparisons (involving sub-samples of 50 to 129 students) there was a significant increase in test performance in the experimental group over the control group, ranging on average from 10 to 25 points.
- In two of these matched-group comparisons, this significant increase in test performance was associated with a significant decrease in test anxiety in the experimental group (Figure 3).

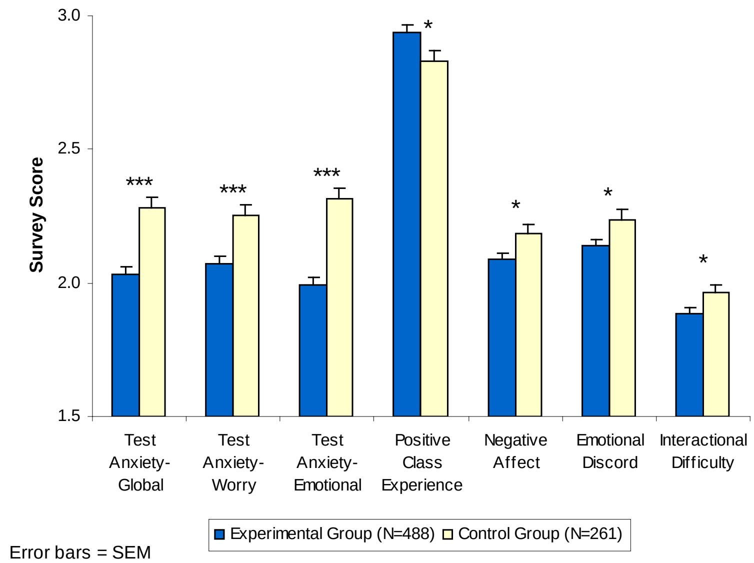


Figure 2. Pre-Post Changes in Test Anxiety and Socioemotional Measures Comparing Intervention and Control Schools
Results of an ANCOVA of pre–post-intervention changes in measures of test anxiety (Global scale, Worry component, and Emotionality component) and socioemotional scales (Positive Class Experience, Negative Affect, Emotional Discord, and Interactional Difficulty) showing significant differences between the intervention and control schools. $*p < 0.05$, $***p < 0.001$.

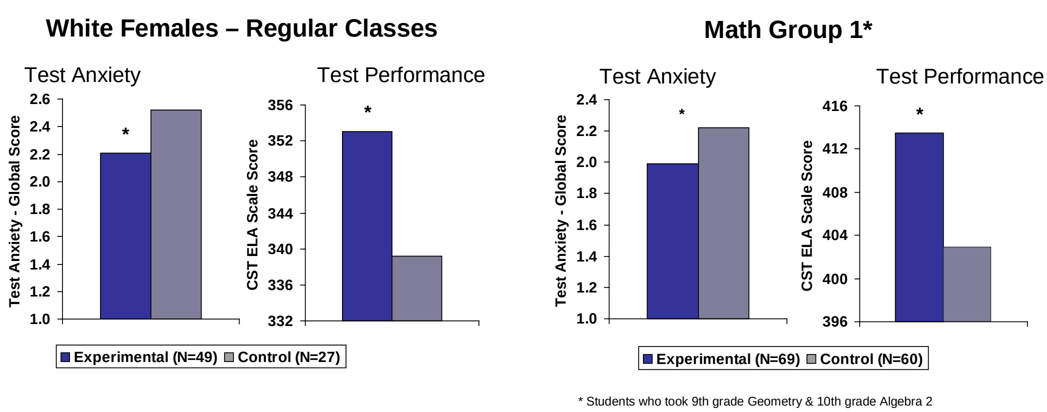


Figure 3. Changes in Test Anxiety and Test Performance in Matched-Group Comparisons
ANCOVA results for two sub-samples from the intervention and control schools matched on sociodemographic factors (White Females in average academic level classes) and 9th grade Math test performance (Math Group 1), respectively. For these matched-group comparisons, significant reductions in test anxiety in conjunction with significant improvements in test performance (California Standards Test – English-Language Arts) were observed in the experimental group as compared to the control group. $*p < 0.05$.

Physiological Study Findings

Results from the electrophysiological study provided compelling evidence for the hypothesis of a causal link between increased psychophysiological coherence and the cognitive functions central to learning and test-taking. In a controlled experiment simulating a stressful testing situation, a random sample of students (N = 136) completed a computerized version of the Stroop color-word conflict test (a standard protocol used to induce psychological stress), while continuous heart rate variability recordings were gathered. For the pre-intervention administration of the experiment, students were instructed to employ whatever methods they typically used when preparing to perform a challenging test or activity. In the post-intervention session, students in the intervention group were instructed to use one of the TestEdge coherence-building techniques they had learned to ready themselves for the test, while the control group students again used their own methods.

The data suggest that when students self-manage their stress using coherence-building methods, it enables them to achieve both a significant reduction in test-related anxiety and a corresponding improvement in standardized test scores. Specifically, results from the post-intervention physiological experiment demonstrated that:

- Students in the experimental group had acquired the ability to self-activate the coherent state prior to taking an important test (see Figure 4 on the next page).
- This ability to self-activate coherence was associated with significant reductions in test anxiety and corresponding improvements in measures of emotional disposition.
- Students in the experimental group also exhibited increased heart rate variability and heart rhythm coherence *during the resting baseline period* in the post-intervention experiment—even without conscious use of the TestEdge tools. This suggests that through their consistent use of the TestEdge tools over the study period, these students had instantiated a healthier, more harmonious, and more adaptive pattern of psychophysiological functioning as a new baseline or norm.

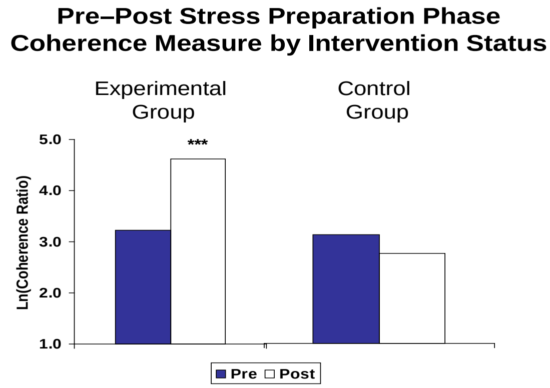


Figure 4. Heart Rhythm Coherence While Preparing for a Stressful Test
These data are from the electrophysiological study—a controlled experiment involving a random stratified sample of students from the intervention and control schools (N = 50 and 48, respectively). In this experiment, students were administered the Stroop stress test while heart rate variability was continuously recorded. These graphs quantify heart rhythm coherence—the key marker of the psychophysiological coherence state—during the stress preparation phase of the protocol. Data are shown from recordings collected before and after the TestEdge intervention. The experimental group demonstrated a significant increase in heart rhythm coherence in the post-intervention recording, when they used one of the TestEdge tools to prepare for the stressful test, as compared to the control group, who used their own stress preparation techniques. ***$p < 0.001$

- In a sub-sample of students matched on baseline test scores, the capacity to self-activate coherence was associated with a reduction in test anxiety as well as an improvement in test scores in the experimental group (see Figure 5 on the next page).

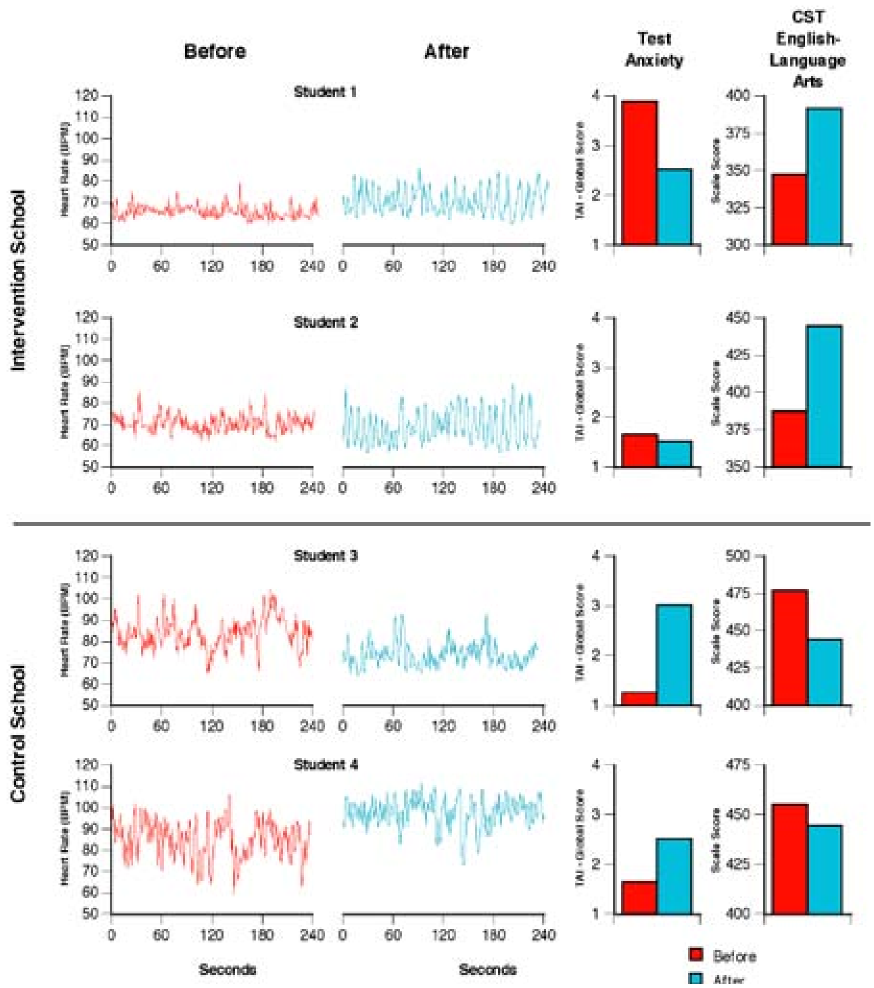


Figure 5. Typical Heart Rate Variability Patterns in Four Students Preparing for a Stressful Test
Heart rate variability (HRV) recordings from the electrophysiological study, showing four students' heart rhythm patterns while they prepared themselves for the Stroop stress test, both before and after the TestEdge intervention. Pre- and post-intervention test anxiety level (TAI-Global Scale score) and the California Standards Test (CST)–English Language Arts test score for each student are also shown. For the two students in the intervention school, the recordings show a shift from an erratic, irregular heart rhythm pattern (left-hand side) before the intervention, to a sustained sine-wave-like pattern (increased heart rhythm coherence), indicative of the coherence state, after the intervention. By contrast, both the pre and post HRV recordings for the students in the control school signify an ongoing incoherent psychophysiological state.

Findings from the Study of Student Drawings

A window into how students see and feel about themselves while taking an important test was provided by an assessment of student drawings collected as a part of the Student Opinion Survey questionnaire administered in the primary study. Altogether, a total of 1,581 drawings were collected (830 pre- and 751 post-intervention). From the analysis of a random sample of 95 pre-intervention drawings, it was found that:

- The overwhelming majority of student drawings conveyed strongly negative feelings and emotions.
- The drawings contained very few depictions of adults which, almost without exception, were negative.

From an analysis of the pre-post pairs of drawings we identified the following patterns of post-intervention change along three dimensions in sub-samples of students (N = 109) drawn from the intervention school:

- Movement from a negative to a positive self-concept.
- Movement from negative to positive feelings and cognitions.
- Movement from negative to positive perceptions of self-control and success.

In contrast to these positive patterns of change, very few instances of positive pre-post change were observed in the student drawings from comparable sub-samples in the control school.

Qualitative Findings

To supplement the quantitative data, the study gathered observations of student classroom interactions in the two schools and conducted structured interviews with teachers.

Classroom Observations

An observational protocol was developed for systematic data collection from three periods of passive observation of the socioemotional environment and interactions patterns in classrooms before, during, and after the TestEdge intervention. The observational findings are broadly consistent with the findings from the quantitative analysis. More specifically, results from a study of pre–post-intervention classroom observations comparing the two schools showed that:

- More positive changes were observed in the classrooms of the experimental school while more negative changes were observed in the control school.
- Students at the experimental school exhibited reduced levels of fear, frustration, and impulsivity. They also exhibited increased engagement in class activities, emotional bonding, humor, persistence, and empathetic listening and understanding.

Teacher Interviews

A number of important findings emerged from the teacher interviews:

- Most teachers acknowledged that their students came to school emotionally unprepared to learn.
- Most teachers felt that their educational training did not equip them with the requisite skills to effectively manage their own stress or to help their students to do so.
- Most teachers were supportive of integrating emotional management instruction into educational curricula.
- Most teachers reported experiencing personal benefits as well as positive changes in their students' behavior as a result of the intervention program.

Major Findings: Secondary Study

Evaluation of the implementation case studies of the TestEdge program, conducted in selected classrooms at various grade levels across different states, produced a number of notable results.

- Corroborating the findings from the primary study, interviews with teachers revealed that:
    - o The lack of emotional self-management education for students was seen as a significant obstacle to learning and academic performance
    - o Few felt they possessed the requisite emotional management skills to effectively manage stress or to teach their students how to do so
    - o Most felt the program provided substantial benefits in both their professional and personal lives
- In relation to the program's impact on their students, the teachers:
    - o Described positive changes in students' attitudes, behaviors, test anxiety and academic performance
    - o Felt that the tools and skills would have a positive impact on students' future socioemotional and academic development

- In general, the program's implementation was more successful when:
    - There were several teachers at the same grade level teaching the course
    - Teachers were able to internalize the use of the tools in their own lives
- In general, major challenges to successful program implementation were:
    - Inadequate class time
    - Logistical, coordination, and communication problems encountered with school administration
    - Securing the support of the principal and other key school administrators to foster teacher commitment

Elementary School Case Study

An exemplary case of a highly successful implementation of the TestEdge program was provided by an in-depth study conducted at the third grade level in a Southern California elementary school. Several notable findings emerged from the study:

- Large increases in state-mandated test scores were observed, which far exceeded academic targets for the year. As a result, student proficiency grew from 26% to 47% in English Language Arts and from 60% to 71% in Mathematics.
- Corresponding emotional and behavioral improvements among students in the classrooms were also observed.
- Success of the implementation was largely due to the enthusiastic support provided by the school's principal and key teachers and administrators.

Conclusion

Overall, the preponderance of evidence from this rich combination of physiological, quantitative, and qualitative data, indicates that the TestEdge program led to a number of important successes. There is good reason to believe that the program produced substantial physiological, psychological, academic, and social benefits for the participating students.

It is our hope that the results of this research will have an impact on policies regarding the importance of integrating stress and emotional self-management education into school curricula for students of all ages. By introducing and sustaining appropriate programs and strategies, it should be possible to significantly reduce the stress and anxiety that impede student performance, undermine teacher–student relationships, and cause physiological and emotional harm. Such programs have the promise of increasing the effectiveness of our educational system and, in the long-term, boosting the academic standing of the United States in the international community.

Acknowledgments[1]

This study was funded by the U.S. Department of Education's Fund for the Improvement of Education, grant number U215K040009.

Oliver Wendell Holmes said, "Man is an omnibus in which his ancestors ride." He might have said that anyone who undertakes a large, complex research study is an omnibus in which numerous people ride. This research study represents a true collaborative effort, one that has enjoyed the imagination, skill, industry, and goodwill of a large number of people. As a research team we acknowledge our indebtedness to all of those friends, colleagues, associates and even distant acquaintances who made contributions large and small to this endeavor. It is our hope that all will feel rewarded for their participation in this study by the expectation that our findings will lead to a reduction in the stress teachers and students experience over testing and thereby give them more time and energy to devote to the truly exciting adventure of learning.

Those who supported and contributed to the project include Congressional representatives, government officials, school administrators, teachers, graduate students, researchers, support staff of many kinds, and, last but not least, the thousands of students who participated in the project.

First and foremost we acknowledge the support of this project by members of Congress and their staffs from both parties. These include Representatives Anna Eshoo, Sam Farr, Ruben Hinojosa, Todd Platts, Thomas Petri, Michael Castle, Bill Young, and Steny Hoyer. Special recognition goes to Representative Ralph Regula, a long-time advocate of improved public education, for his leadership in supporting this and many other educational initiatives.

We would like to give special mention and acknowledgment to the administrators, teachers, and students at the primary study schools. Among these, we express particular appreciation to the following from the intervention school: the school's Superintendent, Vice Principal, Senior English Teacher and Project Coordinator, and Head of Technology. All played key roles in the implementation and data collection process.

[1]To preserve confidentiality, throughout this report the names of the schools and their staff have been replaced with generic or fictitious names.

At the control school, we thank the Principal, Associate Principal, and key teachers who provided logistical and other support in making the implementation a success.

We also acknowledge the support of the many administrators and teachers at the secondary sites. Not only were these educators supportive with their staffs and resources, they were generous and gracious in working with us on the many details required in implementing the program in such diverse schools. We also thank the Principal at the "Southern California" Elementary School, and the many teachers and students for their support and facilitation in the implementation and research process. This school was the focus of an in-depth case study. Without the support and participation of these and many other caring people, this study would not have been possible.

Also of particular note was the participation of a team of highly dedicated and enthusiastic graduate students under of the direction of co-principal investigator Lourdes Arguelles of Claremont Graduate University's School of Education Studies. These include Alane Daugherty, who played many roles, from teacher training, coordinating the observational team and their schedules, to having a leading role in the physiological study. The observational team for the primary sites, who did an exceptional job in designing the protocols and using their vacation time to observe the classrooms on multiple occasions, included Kim Hartnett-Edwards, Teri Hollingsworth, and Laurie Schroeder. Laurie also played a leading role in the design of the observational protocol and the student drawing assessment protocol, in addition to supervising the CGU processing and analysis of student drawings. All deserve special recognition.

The secondary site observer team, who had to contend with long flights, less than optimal accommodations, bad weather, changed schedules, and other vicissitudes, included Nola Butler-Byrd, Avery Caldwell, Jeff Lagozzino, Manuel Rodriguez, and Emily Wolk. Andy Behr, Sam Bauman, Peter Chen, and Adonay Montes, also from CGU, were involved in data processing and analysis.

We would also like to thank Dr. David Drew, Professor of Education (joint appointments in Executive Management, Psychology, and Mathematical Sciences) at CGU, for his enthusiastic support and dedication to the project—in particular for his expert input to our statistical analysis plan and for his detailed review and helpful comments on an earlier draft of the quantitative analysis results.

We owe many thanks to staff at the Institute of HeartMath and at the HeartMath Research Center. Christiana Bishop and Jeff Goelitz were the primary coordinators of training logistics and interfacing with the schools. Jeff and Christiana also conducted the Resilient Educator and TestEdge trainings with the teachers at all the school sites in both the primary and secondary studies. Also playing important supportive roles were Brian Kabaker, who oversaw the financial administration of the grant, and Lynn Lippold, who dealt with all the tracking and auditing aspects of the project. Del Walker played a valuable role as a member of the classroom observation team, and Glenda Albert helpfully assisted with data collection.

From the HeartMath Research Center, Jackie Waterman deserves special mention. She was involved in every aspect of the study, from administration, data collection, and scanning and double-checking of the surveys to the set-up of the physiological study and the scheduling and arrangements for all travel.

We are also grateful to Doc Childre, founder of the Institute of HeartMath and creator of the HeartMath system, for his wisdom and care in designing the tools and techniques taught in the Resilient Educator and TestEdge programs.

Finally, we express our gratitude to the many teachers who participated in the Resilient Educator and TestEdge trainings and who made room for the HeartMath program in their classes. We especially appreciate the care and skill with which they introduced the program to their students. Ultimate appreciation goes to the thousands of students who participated in this study for the role they played in helping other students live more stress-free and therefore more happy and productive lives. We wish all of them well on their continuing journeys through school and life. Our experience with them affirms our belief that their wisdom, energy, and talent are vital to the future of the nation.

Table of Contents

Chapter X

Chapter XI

Section
One

Research Background
and
Study Design

Chapter I
Background to the TestEdge National Demonstration Study

With the passage of the No Child Left Behind legislation in 2002 and with the increasing number of state-mandated tests, student test performance has risen to the top of both local and national education agendas. Whatever the ultimate outcome of the debate on the utility of standardized testing, it is clear that testing continues to be an integral part of the American education experience.

The increased emphasis on standardized testing has led to increased awareness of and concern about the test-related anxiety experienced by many students. Research shows that the emotional turmoil caused by test anxiety prevents students from performing up to their true ability on standardized tests. Moreover, administrators and teachers are under increasing pressure to make "adequate yearly progress," which often further increases students' test-related emotional stress. This is particularly true of urban schools with large minority student populations. In some ways this stress, added to the normal stresses and demands characteristic of today's educational environment, is creating a national crisis.

Within this context, the challenge for educators is how best to prepare students for testing so that their performance is reflective of their true ability, and, so that test results provide educators with accurate performance data necessary for curricular and administrative decisions. In order to accomplish this goal, it is necessary to deepen our understanding of student test anxiety and to implement effective tools to help students manage this particular kind of emotional stress as well as the myriad of other stressors in their lives.

The TestEdge program is a supplemental learning program designed to help students reduce stress and test anxiety, improve test and academic performance, and enhance emotional and relational competence. Based on 15 years of research on the psychophysiology of learning and performance, emotional dynamics and heart–brain communication at the Institute of HeartMath and elsewhere, this program teaches a set of easy-to-use, positive emotion-based tools and techniques that enable students to recognize and self-regulate test-related anxiety and other emotionally-based blocks to learning. As described in greater detail below, the basis of the effectiveness of the TestEdge techniques is that they enable the individual to self-activate a specific, scientifically measurable psychophysiological state—*psychophysiological coherence*—associated with increased synchronization in nervous system activity, increased emotional stability, and improved cognitive performance (McCraty, Atkinson, Tomasino, & Bradley, 2006). The TestEdge program has been successfully implemented in schools throughout the U.S. and in some foreign countries, and pilot studies in several high schools have measured associated improvements in standardized test scores and passing rates, as well as global improvements in psychosocial functioning (Arguelles, McCraty, & Rees, 2003).

In 2004, a Department of Education-funded national study was initiated to determine the impact of this learning program on a large student population. This study's primary objectives were: (1) to determine the nature, magnitude, correlates, and consequences of test anxiety among high school students; and (2) to investigate the efficacy of the TestEdge program in reducing test anxiety and improving performance on standardized tests. More broadly, the study also investigated the association between exposure to the program and a range of socioemotional variables, including stress, emotional well-being, attitudes toward school, quality of relationships, and classroom dynamics.

Since an educational curriculum is only as effective as its implementation, a secondary objective of the study was to investigate the programmatic aspects of implementing TestEdge in school systems with diverse characteristics and student populations.

Meeting these broad objectives required us to conduct two studies with different objectives and designs. The primary study utilized a quasi-experimental, pre-post longitudinal research design with matched, randomly assigned experimental and control groups. Employing both quantitative and qualitative measures within a multi-methods framework, this study aimed to rigorously investigate the relationship between the intervention and resultant changes in student psychophysiological states, attitudes, aspirations, social and classroom behaviors, test anxiety, and test performance. Two high schools in Northern California participated in the primary study.

The secondary study was designed as a qualitative case study investigation of the accessibility, receptivity, coordination, and administration of the TestEdge program in a wide variety of school systems with different sociocultural, administrative, and situational characteristics. This study consisted of a series of qualitative case study investigations of schools in eight states (California, Delaware, Florida, Ohio, Maryland, Texas, Wisconsin, and Pennsylvania). These investigations were designed as separate case studies in which the schools were treated as convenience samples. Although rigorous generalizability is not required for this component, the observational and interview data from this study provide valuable information on best practices and pitfalls when implementing interventions such as TestEdge in schools with widely diverse compositions and characteristics.

The results of these two studies have a number of important potential policy implications at the state and national levels. By advancing our understanding of the factors affecting student learning and test performance and introducing appropriate programs and strategies, it should be possible to significantly reduce the stress and anxiety that impede teacher and student performance. In turn, this should increase the effectiveness of our educational system and, in the long-term, help to boost the academic standing of the United States in the international community (Drew, 1996).

Chapter II
Test Anxiety: Its Emotional and Psychophysiological Basis and Relational Context

Many contemporary scientists believe that the quality of feeling and emotion we experience in each moment is rooted in the underlying state of our physiological processes. This was the essence of the theory of emotion first proposed by William James more than a century ago (1884), which has since undergone much further development. Recent expression of this view is expounded by neuroscientist Antonio Damasio (2003: 131): "The fact that we, sentient and sophisticated creatures, call certain feelings positive and other feelings negative is directly related to the fluidity or strain of the life process." The feelings we experience as "negative" are indicative of body states in which "life processes struggle for balance and can even be chaotically out of control." By contrast, the feelings we experience as "positive" actually reflect body states in which "the regulation of life processes becomes efficient, or even optimal, free-flowing and easy" (Damasio, 2003: 131).

Introduction

This is the broader perspective that informs the understanding we develop in this chapter—of test anxiety as fundamentally rooted in the emotional processing of the body's response to both external and internal stimuli. In these terms, the experience of test anxiety is a struggle to maintain emotional composure and self-control in the face of a threatening life space. Thus, when a student faces an upcoming important test, the feelings and emotions activated by the preparation, the anticipation, the pressure of expectations, the test-taking experience, the unknown outcome and consequences, all combine to create a challenging situation fraught with difficulty, uncertainty, fear, and even dread. In turn, these "negative" feelings and emotions are reinforced by the body's experience of the physiological responses activated during this state: increased heart rate, upset stomach, rise in adrenaline and other stress hormones, sleeplessness, fatigue, irritability, and so forth. Through this reciprocal interaction this "strained" psychophysiological state, experienced generally as anxiety, is generated, amplified, and maintained, with deleterious consequences for learning and academic performance.

If, on the other hand, the student is able to activate a positive emotion and thereby intentionally shift his/her bodily state to one of increased "fluidity"—synchronization and harmony in the activity of the body's somatosensory systems and subsystems—this changed psychophysiological milieu acts to generate, reinforce, and stabilize a "positive" feeling state in which the body's regulation of life experience becomes efficient, or even optimal, free-flowing and easy" (Damasio, 2003: 131). As we will see in Chapter III, the result of this shift is an optimal psychophysiological state for learning and academic performance.

Following a brief examination of the sources and prevalence of student stress and test anxiety, in this chapter we develop a psychophysiological perspective which elaborates these ideas. This perspective emphasizes the key role of emotions in both understanding and managing student anxiety and test stress. As described in the review of research that follows, emotion has been found to be an omnipresent and dominant force in driving the body's physiological responses, motivating behavior, and influencing cognitive function. Indeed, the importance of emotions in creating the requisite context for successful education cannot be overstated. In a new paper, Immordino-Yang and Damasio (2007: 3) point to the important connection between emotion, social functioning, and decision-making established by recent work in neuroscience. Moreover, they go on to note that on the basis of the neurological evidence, the elements of cognition targeted in schools—"learning, attention, memory, decision making, and social functioning"—are "both profoundly affected by and subsumed within the processes of emotion."

This conclusion is in accord with the body of evidence generated from the Institute of HeartMath's own research over the last 15 years. Moreover, from an applied perspective, our research has shown that specific positive emotion-focused tools and techniques, which are the foundation of the TestEdge program (described in the next chapter), facilitate an intentional shift to psychophysiological coherence. By doing so, these tools are not only effective in helping students manage stress and overcome other emotional impediments to education, but, by facilitating the shift to an optimal psychophysiological state, they also promote learning, academic performance, and social functioning.

Sources of Student Stress

A student's learning and performance are continuously impacted by the interactions among a complex array of physiological, mental, emotional, sociocultural, and relational factors. Some of the more well-recognized influences include students' level of engagement in the learning process, their motivation to learn and perform, and their attitudes toward school, teachers, and subject matter. As Daugherty (2006: 17) notes, it is almost impossible to separate students' experience in school from the stresses and anxieties of their lives in general, which also exert a significant influence on school-related attitudes, behavior, and performance. Thus, to understand students in a truly holistic sense, it is necessary to consider the general stresses of being a child or adolescent in today's world. These include a host of day-to-day worries, concerns, hurts, misunderstandings, and conflicts in relationships with peers and family; changes related to physiological maturation; peer pressures to conform to social norms and/or to engage in risky behaviors; the increasing allure of media, technology and the Internet; awareness of local and global conflict and change (war, crime, natural disasters, environmental changes, etc.); the intense demands of school itself as well as full schedules of after-school activities; and increasing parental and societal expectations of adult-like competence in children, a trend that has been referred to as the "superkid syndrome" (Elkind, 2001).

Youth who come from broken homes or single-parent families face the added stresses associated with parental separation or divorce and often increased household responsibilities, such as working and caring for younger siblings. Beyond these challenges, many children and adolescents in the U.S. live in situations where the stresses of everyday living are overwhelming and violence and poverty rampant. Also, like adults, children are impacted by significant global threats such as terrorism, war, and the increasing number of natural disasters caused by climate change. When the feelings and emotions resulting from such diverse stressors are added to the array of physiological, emotional, and relational changes that accompany the transition from childhood and adolescence, it is clear that the stress on young people's daily lives can have a significant, and for some, a debilitating impact (Daugherty, 2006).

In addition to the age-related stresses of childhood and adolescence, and the stresses of achievement expectation, violence, poverty, and global threats, a student's experience at school carries an additional burden of anxieties and fears. Anxiety about school in general, conflicts with teachers, difficulty with classmates, performance pressures, learning disorders, academic failure, bullying, and teasing are but a few common school-related stressors (Sears & Milburn, 1990). For the marginalized child or adolescent—that is, those additionally challenged by social or psychological pressures beyond the norm—the weight of such fears and anxieties may grow to overwhelming proportions (Daugherty, 2006: 18).

Prevalence of Test Anxiety

Over and above their general anxiety about school, students also experience a significant amount of anxiety specifically related to taking tests. This anxiety is manifest mentally, emotionally, and physically. Anxiety in its state, trait, and psychopathological forms is grounded in the emotion of fear (Barlow, 1988). Because test anxiety is often associated with other anxieties and phobias, it is difficult to provide a simple, clear-cut definition. Most researchers view test anxiety as a form of state anxiety that falls in the broader category of performance anxiety. State anxiety is experienced in different specific situations and to different degrees. An important distinction between trait anxiety and state anxiety is that a student who experiences state anxiety in a specific situation (e.g., when speaking in front of the class, performing in a concert or on an athletic field, or taking a test) can be relatively low in trait anxiety. However, a student with higher trait anxiety—meaning he or she is more generally anxious—is likely to be more anxious across a variety of specific situations, including high-stakes testing (Spielberger & Vagg, 1995: 6-7).

One of the more familiar definitions of test anxiety was provided by Zeidner (1998). He defines test anxiety as:

> *…the set of phenomenological, physiological, and behavioral responses that accompany concern about possible negative consequences or failure on an exam or similar evaluative situation… . Test anxious students are characterized by a particularly low threshold for anxiety in evaluative situations, tending to view evaluative situations, in general, and test situations in particular as personally threatening… . Test-anxious behavior is typically evoked when a person believes that her or his intellectual, motivational, and social capabilities and capacities are taxed or exceeded by demands stemming from the test situation* (pp. 17-18).

Beyond the general psychological and physiological impact of test anxiety, of growing concern is the effects of anxiety on test performance, since students with moderate to high levels of test anxiety have been shown to have depressed test performance (Hembree, 1988). In 1984, K. T. Hill estimated that "as many as 10 million students in elementary and secondary schools perform more poorly on tests than they should because anxieties and deficiencies in test-taking strategies interfere with performance" (quoted in Wigfield & Eccles, 1989, p. 159). It should be emphasized that this estimate of 10 million was made in 1984—22 years ago. With the growth in student population and the exponential increase of standardized testing in recent years, it is likely that this figure has risen dramatically (Daugherty, 2006). Thus, test anxiety represents an important testing artifact that detracts from accurate measurement of a student's true level of academic knowledge or skill.

In previous studies, estimates of the percentage of students who suffer from test anxiety have ranged from approximately 1% to 40% (Cizek & Burg, 2006). Such wide variability is likely due to the different populations studied and to differences in the instrumentation and operational definitions employed. This makes the task of interpreting the meaning of such widely varying estimates extremely difficult. For example, a study giving an estimate of 1.1% refers to the percentage of students who worry about making mistakes (Beidel, 1991), whereas another study that estimated 10% refers to elementary students who are highly anxious (Erford & Moore-Thomas, 2004). Higher estimates of test anxiety were provided by studies which find that 25% to 30% of children suffer from "debilitating stress" in evaluative situations (Hill, 1984), while 41% of African-American students in the third through sixth grades are significantly affected by test anxiety (Turner, Beidel, Hughes, & Turner, 1993). This paucity of data on the general incidence of test anxiety does not appear to have improved much in the last decade or so. Writing in a recently published book on test anxiety, Cizek and Burg (2006: 29) quote from Zeidner's authoritative 1998 book, *Test Anxiety: The State of the Art*, that " 'the data on the prevalence and incidence of test anxiety are surprisingly sparse' (Zeidner, 1998: 6)."

It should now be apparent that most estimates of test anxiety come from studies conducted prior to the enactment of the No Child Left Behind Act (NCLB) of 2002, which many believe has been a catalyst for a major increase in stress and anxiety levels in administrators, teachers, and especially students. The NCLB Act has mandated annual testing in reading, mathematics, and science for every student in grades 3 through 8 and at one point during high school, as well as monitoring of yearly progress for all schools. One of the goals of our study was to provide current data on test anxiety in America's schools by measuring test stress in two large high schools in California using an established, sound measurement instrument: Spielberger's *Test Anxiety Inventory* (Spielberger, 1980)

The accountability for student test performance falls on teachers and administrators, who are expected to produce annual gains in test scores. For many educators, this has led to increased stress and anxiety and dissatisfaction with the educational system. Since previous research shows a strong correlation between students' test anxiety and their teachers' general anxiety, this increase in teacher stress is likely an important factor in increasing student test anxiety and, in turn, depressing test performance.

In addition to the NCLB-imposed tests, most states have mandated that students pass an "exit examination" prior to being awarded a high school diploma. Typically, students are required to pass a series of tests in core subjects such as English-language arts, mathematics, writing, and science. These tests directly impact students by determining whether and when they graduate from high school. In many cases these outcomes affect students' future options and prospects. As such, they have significant economic and social implications for American society.

A Cognitive Model of Test Anxiety

Most research on test anxiety, and, correspondingly, most interventions for addressing it, adopt a primarily cognitive perspective—giving primacy to the cognitive processes that influence the anxiety response. Building upon Lazarus's (1966) conception of stress as a "transactional process," Spielberger (1966, 1976) developed a model of test anxiety which distinguished between the stress associated with testing situations (the stressor), the subjective evaluation of the degree of threat a given test poses to the individual (the threat), and the emotional state of anxiety evoked in the individual in response to the perceived threat—feelings of tension, apprehension, nervousness, and worry, including the associated physiological arousal generated by activation of the autonomic nervous system. The premise of the Spielberger model is that the intensity of the anxiety reaction "will vary as a function of the degree of perceived threat" (Spielberger & Vagg, 1995a: 6). In other words, the anxiety reaction is driven by a cognitive evaluation of the perceived potential threat posed by a test. In short, the model views test anxiety as the outcome of a specific temporal sequence of events (Spielberger & Vagg, 1995a: 6-7):

Stressor → Threat → Anxiety

According to Spielberger's model, the process starts when a person is faced with a challenging task, such as a test. The second step in the process is the formation of cognitive perceptions about the difficulty of the task. These perceptions are affected by the amount of preparation the individual has undertaken (knowledge and study skills) and his or her perceived test-taking skills. The next step involves the subconscious internal evaluation of the accuracy of these perceptions. This evaluation is ongoing and cyclical as the student constantly reforms perceptions about self and the task and appraises the accuracy of these perceptions. For students with relatively higher levels of trait anxiety, the internal perceptions and appraisals result in a view of the challenge as threatening. This, in turn, increases physical/autonomic stress responses (termed "emotionality" in some test anxiety research) and worry, which interfere with cognitive processes. These effects influence the last step in the process, the response produced in testing situations. The response can facilitate or inhibit test performance, thus affecting the test's ability to measure the student's true level of knowledge or skill (Spielberger & Vagg, 1995a; Cizek & Burg, 2006).

In the model of test anxiety just described, it is the *cognitive perception and appraisal* of a challenge (the test) as threatening that is viewed as driving the consequent activation of a set of physiological, psychological, and behavioral reactions which can be characterized as the "anxiety response." According to the cognitive perspective, all *emotional* aspects of the anxiety response necessarily follow a cognitive assessment of the stressor; it is therefore presumed that by changing one's thoughts about a potentially threatening stimulus, one can gain control over one's emotions. However, this presumption belies the enormous, omnipresent influence that emotions are now known to have on virtually all aspects of cognition and behavior, as we will see in the rest of this chapter.

Emotional Basis of Anxiety

Recent research in the neurosciences has significantly broadened our understanding of the workings of the emotional system itself, as well as its extensive interactions with cognitive function. On the basis of this new understanding, emotion and cognition can best be thought of as separate but interacting functions and systems which communicate via bidirectional neural connections between the neocortex and emotional centers such as the amygdala. These connections allow emotion-related input to modulate cortical activity and cognitive input from the cortex to modulate emotional processing.

However, research has shown that *the neural connections that transmit information from the emotional centers to the cognitive centers in the brain are stronger and more numerous than those that convey information from the cognitive to the emotional centers* (LeDoux, 1996). This accounts for the powerful influence of input from the emotional system on virtually all stages of cognitive processing involved in functions such as attention, perception, and memory, as well as on higher-order thought processes, like logical reasoning and rational decision making. This fundamental asymmetry also provides a physiological basis for the common experience that emotions such as anxiety can readily dominate the mental landscape, yet it is usually far more difficult to willfully "turn off" these strong feelings through thought alone.

Moreover, it is now clear that the emotional system can also operate entirely independently of the cognitive system. For example, studies have found that emotional processes operate at a much higher speed than thoughts and frequently bypass the mind's linear reasoning process entirely (LeDoux, 1996). This has been described in more popular terms as "emotional hijacking" (Goleman, 1995). In other words, not all emotions follow thoughts; *emotions often occur without involvement of the cogni-*

tive system and, moreover, can significantly color the cognitive process and its output (LeDoux, 1996; LeDoux, 1994; Niedenthal & Kitayama, 1994).

Such is the case when emotional memories of past threatening experiences automatically trigger a fear-anxiety response to a future anticipated event, often circumventing the processes of conscious thought and self-control. LeDoux's work (1996) provides an understanding of the mechanisms involved. Evolving long before the neocortex, the subcortical brain circuitry involved in emotional processing is highly attuned to signs of potential danger and it is hyperreactive to perceived threat. Through a process called fear conditioning, the body can learn to perceive an otherwise mundane stimulus as threatening. The amygdala forms a key part of this subcortical circuitry, and it is responsible for processing subconscious emotional memory in which it plays a significant role in the activation of fear. Even before the cortex is able to consciously perceive and respond to a threat, the amygdala has already activated the body's stress response, causing a flood of biochemical and cardiovascular reactions.

Within this context of test anxiety, the important point is that emotional memories can be triggered by the anticipation of a future event that is *similar* to a past unpleasant event, *irrespective of whether or not those emotions are appropriate for the current situation* (LeDoux, 1996). In each new situation, the amygdala takes in sensory input across the full range of bodily experience (sights, sounds, smells, facial expressions, perceptions of nonverbal behavior, etc.) and, through a pattern-matching process (described in a later section), looks for a match between these current sensory inputs and those stored as past emotional memories. Once the amygdala finds a match or near match, it triggers a system-wide physiological and psychological response.

Taken as a whole, these new understandings of the extensive interactions between the emotional and cognitive systems have enormous implications for education, as Immordino-Yang and Damasio (2007) elaborate in a recent paper aptly titled "We Feel, Therefore We Learn." As noted at the outset of this chapter, they conclude that the very "aspects of cognition that we recruit most heavily in schools, namely, learning, attention, memory, decision making, and social functioning, are all profoundly affected by and subsumed within the processes of emotion" (Immordino-Yang & Damasio, 2007: 3). This influence of emotion on cognition is clearly apparent in their model shown in Figure II.1.

In short, this perspective helps explain why it is most often one's feelings and emotions, rather than thoughts and cognitions alone, that are the most powerful drivers of physiological responses and the strongest motivators of behavior. For this reason, interventions that focus solely on mental processes may often fail to identify the fundamental source of an emotional disturbance such as anxiety, and thus to resolve it. In some cases, try as one might to rectify one's thinking, one can fall short of achieving emotional relief simply because the underlying maladaptive emotional pattern are driven largely by unconscious processes that operate independently of the intellect.

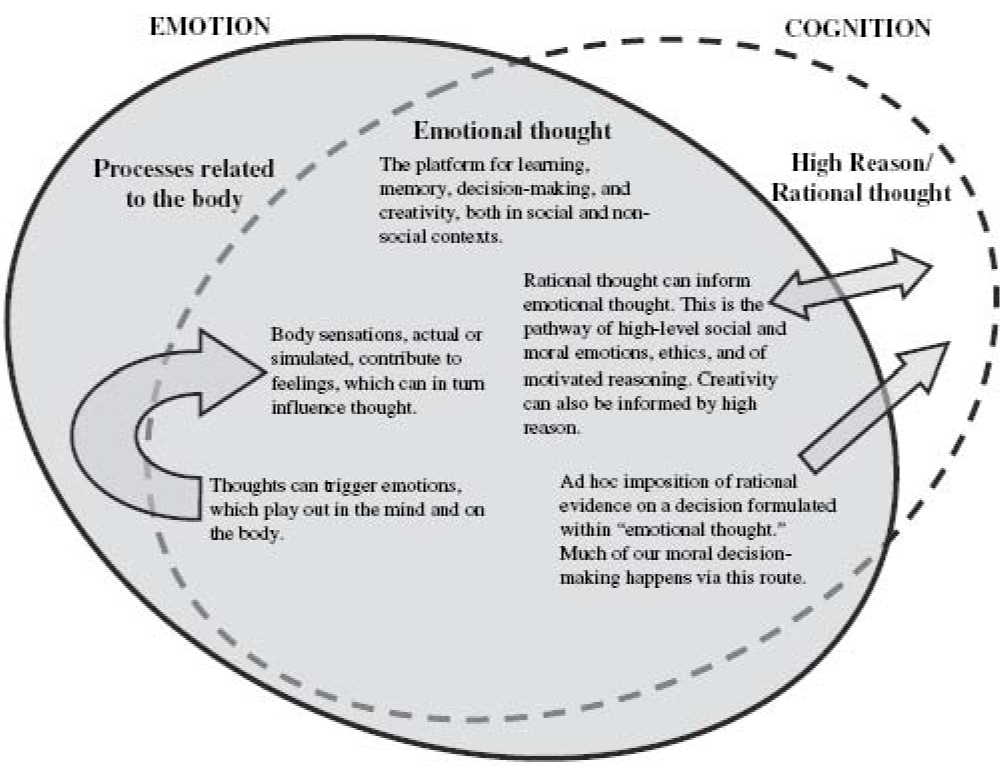


Figure II.1. A Model of the Influence of Emotions on Cognition.
"The evolutionary shadow cast by emotion over cognition influences the modern mind. In the diagram, the solid ellipse represents emotion; the dashed ellipse represents cognition. The extensive overlap between the two ellipses represents the domain of emotional thought. Emotional thought can be conscious or nonconscious and is the means by which bodily sensations come into our conscious awareness. High reason is a small section of the diagram and requires consciousness" (from: Immordino-Yang & Damasio, 2007, Figure 1, page 8; reproduced with permission).

Psychophysiology of Anxiety

Understanding the psychophysiological manifestations and effects of anxiety is crucial in comprehending its tremendous impact on student cognition, learning, and academic performance. The somatic expression of anxiety is often portrayed as a preparation to flee, in accord with Cannon's (1929) "flight-or-fight" model of response to threat. Cannon emphasized increases in sympathetic nervous system activity to optimize blood flow and metabolism, as reflected in cardiovascular changes such as increased heart rate and blood pressure. Hence, it is reasonable to expect that anxiety states, by way of their close relationship to fear, would be associated with autonomic nervous system (ANS) activation; this notion is indeed well supported. In addition to the changes in ANS activity, research has also found that habitual increases in stress hormones, such as cortisol, can produce an increase in brain receptor sites for these chemicals. This increases the physiological likelihood of perpetuating and amplifying these stress-induced states (Rosenzweig, Leiman, & Breedlove, 1999). Research has also shown that routine low-level "cortisol baths" significantly contribute to the development and onset of depression (National Institute of Mental Health, 2000).

When a challenge is perceived as threatening and triggers feelings of fear and anxiety, this response can manifest in a wide range of symptoms, including sleep disturbance, withdrawal, vomiting, sweating, crying, throwing tantrums or wetting themselves (in younger children), inappropriate behaviors, cheating, or sudden illness, especially the night before a test (Edelstein, 2000: 2). Longer-term effects of test anxiety include erosion of academic motivation and positive attitudes towards education and learning, and reinforcement of negative self-perceptions of confidence and ability to learn (Cizek & Burg, 2006).

One of the often underappreciated effects of anxiety is that it distracts attention. The importance of this is clear from research in neuroscience that has shown that we only perceive what we attend to—a phenomenon termed *inattentional blindness* (Most, Scholl, Clifford, & Simons, 2005). The implication of this is that when a student's attention is sufficiently distracted by intense feelings of anxiety, it can severely compromise his or her ability to fully attend to, and thus correctly perceive, comprehend, and respond to the information on a test. In some cases, anxiety-induced inattentional blindness may cause a student to misread test questions and therefore answer incorrectly, as a result of not consciously "seeing" and mentally processing critical words, phrases, or other visual information contained in the questions.

Anxiety and worry also generate the equivalent of mental "noise" in the brain, overloading the neural circuits that are otherwise available for and involved in higher-order cognitive processes. Research has shown that the psychophysiological activity associated with heightened anxiety and other negative emotions interferes with the brain's ability to properly synchronize neural activity (Ratey, 2001). The resulting de-synchronization inhibits brain processes necessary for functions such as attention, memory recall, abstract reasoning, problem solving, and creativity. Thus, when students come to school or enter a testing situation with high levels of anxiety and emotional stress, the resulting "inner noise" impairs the very cognitive resources needed for learning, memory, and effective academic performance (Arguelles, McCraty, & Rees, 2003; McCraty, 2005).

Emotions and Heart Rhythm Patterns

Research at the Institute of HeartMath has shown that the physiological desynchronization associated with anxiety and other negative emotions is also reflected in patterns of heart activity. This research utilized an important measure of the heart's rhythmic activity called heart rate variability. This is described briefly in what follows.

Rather than beating at a constant rate and thus generating a steady rhythm, the beat-to-beat activity of a healthy heart under resting conditions is actually quite irregular, reflecting the nervous system's variable adaptiveness variability to inputs from both inside and outside the body. This natural beat-to-beat fluctuation in heart rate, generated by the dynamic interplay of many of the body's systems, is known as *heart rate variability* (HRV; see Figure II.2). Short-term (beat-to beat) changes in heart rate are largely generated and amplified by the interaction between the heart and brain, mediated via the flow of neural signals through the efferent (descending) and afferent (ascending) pathways of the sympathetic and parasympathetic branches of the ANS. Scientifically, HRV is thus regarded as a measure of neurocardiac function that reflects heart–brain interactions and autonomic nervous system dynamics. When beat-to-beat changes in heart rate are plotted over time, the overall shape of the waveform produced is called the *heart rhythm pattern* (examples are shown in Figure II.3).

Emotions such as anxiety, fear, anger, and frustration produce heart rhythm patterns that appear *incoherent*—disordered and erratic (McCraty et al., 2006; Tiller, McCraty, & Atkinson, 1996); see Figure II.3. Studies have shown that prefrontal cortex activity affects patterns of heart activity via modulation of the parasympathetic branch of the ANS (Lane, Reiman, Ahern, & Thayer, 2001); therefore, disordered activity in higher-level brain systems manifests as increased disorder in heart rhythm patterns. Physiologically, incoherent heart rhythm patterns are also indicative of desynchronization in the reciprocal action of the parasympathetic and sympathetic branches of the ANS (McCraty et al., 1995; Tiller et al., 1996). This ANS desynchronization taxes the nervous system and bodily organs and thus impedes the efficient flow of information throughout the psychophysiological systems (McCraty et al., 2006).

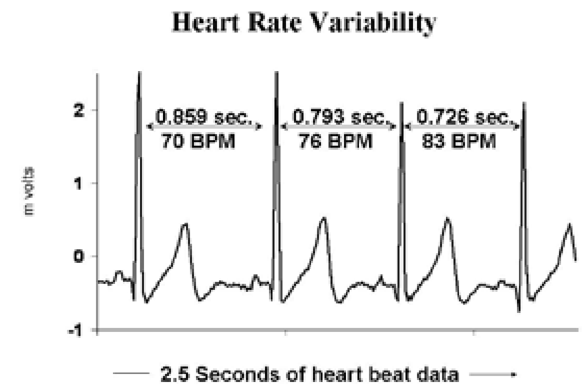


Figure II.2. Segments of ECG recording showing beat-to-beat variability in resting heart rhythm. This diagram shows three heartbeats recorded on an electrocardiogram (ECG). Note that variation in the time interval between consecutive heartbeats, yielding a different heart rate (in beats per minute) for each interbeat interval. This natural beat-to-beat variation in heart rate is known as *heart rate variability* (HRV).

In contrast, sustained positive emotions, such as appreciation, care, compassion, and love, generate a smooth, ordered, sine-wave-like pattern in the heart's rhythms. This pattern reflects increased synchronization between the two branches of the ANS and a general shift in autonomic balance towards increased parasympathetic activity. As is visually evident (Figure II.3) and also demonstrable by quantitative methods (Tiller et al, 1996; McCraty et al, 2006), heart rhythms associated with positive emotions, such as appreciation, are clearly more *coherent*—organized as a stable pattern of repeating sine waves—than those generated during a negative emotional experience such as anxiety.

Figure II.3

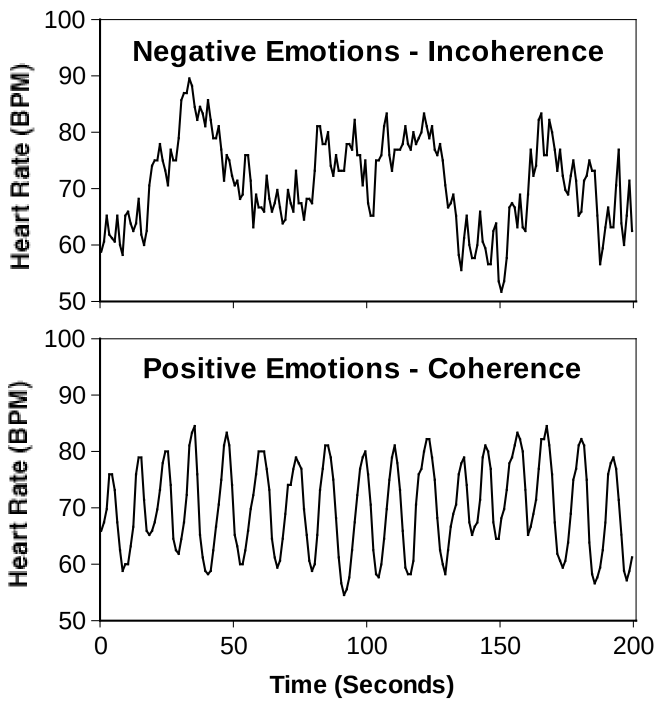


Figure II.3. Heart rhythm patterns reflect different emotional states.
These heart rate tachograms show examples of the heart rate variability patterns recorded in real time from individuals experiencing different emotions. Negative emotions, such as anxiety, anger, and frustration, typically give rise to an erratic, irregular heart rhythm pattern (incoherence). In contrast, positive emotions, such as appreciation, care, and compassion, produce a highly ordered, stable heart rhythm pattern of smooth, repeating sine waves (coherence).

Heart Activity Affects Brain Function

An important, but lesser known consideration, is that patterns of heart rhythm activity not only *reflect* brain processes involved in thought and emotion, but also *affect* these processes via the flow of neural signals through the cardiovascular afferent (ascending) nerves to the brain. The effect of heart activity on brain function has been researched extensively over the last half century. It is now known that the heart actually sends more neurological signals to the brain than the brain sends to the heart (Cameron, 2002). Moreover, it has been shown that these heart signals have a significant effect on brain function—not only exerting homeostatic effects via their interaction with cardiovascular and respiratory regulatory centers in the brain, but also influencing the activity and function of higher brain centers involved in perceptual, cognitive, and emotional processing (see McCraty et al., 2006, for a review).

Research has also demonstrated that different patterns of cardiac activity have distinct effects on cognitive and emotional function. For example (see Figure 11.4), during emotional stress such as anxiety, when patterns of heart activity are erratic and disordered, the corresponding patterns of neurological signals traveling from the heart to the brain produce an *inhibition* of higher cognitive functions. This limits one's ability to think clearly, focus, remember, learn, and reason. The heart's input to the brain during stressful or negative emotions also has a profound effect on the brain's emotional processes—both compromising emotion regulation and reinforcing the emotional experience of stress.

In contrast, the more ordered and stable pattern of the heart's input to the brain during positive emotions has the opposite effect—serving to *facilitate* cognitive function and reinforcing positive feelings and emotional stability. This is a particularly important point in understanding the operative mechanism of the HeartMath techniques taught in the TestEdge program, and will be elaborated in Chapter III.

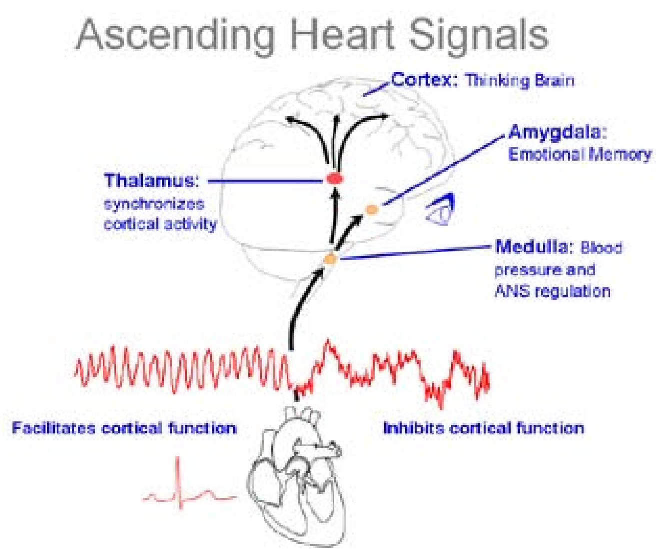


Figure II.4. Heart activity affects brain function.
This diagram illustrates afferent (ascending) pathways by which neurological signals generated by the heart are transmitted to key centers in the brain. These heart signals not only impact autonomic regulatory centers in the brain (e.g., the medulla), but also cascade up to higher brain centers involved in emotional and cognitive processing, including the thalamus, amygdala, and cortex. By these pathways, heart activity exerts a continuous impact on numerous aspects of brain function. As shown, when patterns of heart activity are erratic and disordered, such as during emotional stress, the corresponding patterns of neurological signals traveling from the heart to the brain produce an *inhibition* of higher cognitive and emotional functions. In contrast, the more ordered and stable pattern of the heart's input to the brain during positive emotions has the opposite effect—serving to *facilitate* cognitive function and reinforcing positive feelings and emotional stability.

Significance of the Parasympathetic Nervous System

As is readily apparent, there are multiple physiological pathways involved in test anxiety. Peripheral sympathetic nervous system activity, as well as its central nervous system control architecture, and hormonal and neurotransmitter activity, all play a significant role. However, there has been a tendency in most anxiety research to emphasize the sympathetic (activating) branch of the ANS and to neglect parasympathetic (inhibitory) function. Metaphorically speaking, researchers were searching for a "sticky accelerator" while overlooking the possibility of "bad brakes" (Friedman, in preparation). This approach was reversed in the 1990s, which saw an upsurge in research that utilized HRV analysis as a window into autonomic nervous system dynamics and heart–brain interactions.

Over the last decade, psychophysiologists Bruce Friedman and Julian Thayer have developed a model that relates anxiety to parasympathetic nervous system activity, autonomic flexibility, and adaptivity (Friedman & Thayer, 1998a, 1998b; Thayer & Friedman, 1993, 1997). Their approach reconsiders the traditional homeostatic models and emphasizes the importance of inhibitory processes in the self-regulation of emotions such as anxiety. Their model draws from findings that show that anxiety in all its forms is associated with aberrant autonomic nervous system control of the heart. Multiple HRV indices implicate low parasympathetic activity and elevated sympathetic activity in anxiety. This supports a view of anxiety as a systemic inflexibility grounded in poor inhibitory control.

Low parasympathetic activity has been linked with poor emotion regulation, decreased reactivity to various stimuli, and increased stress vulnerability in infancy and childhood (Porges, 1992b; Porges, Doussard-Roosevelt, & Maiti, 1994), delinquency risk in preadolescent boys (Pine et al., 1998), and both anxiety and antisocial behavior in adolescents (Mezzacappa et al., 1997). In addition, HRV differences may emerge at an early developmental stage in anxiety disorders. Youth with anxiety disorders show higher resting heart rate and a blunted HRV response to novelty, suggestive of impaired vagally mediated parasympathetic cardiac control (Monk et al., 2001). In contrast, increased vagal tone has been associated with increased physiological and behavioral flexibility, responsiveness to the environment, stress resiliency, and emotion regulation ability (Porges, 1992b; Porges et al., 1994). Evidence suggests that high vagal tone also enhances the cognitive processes involved in learning, including attentional capacity

(Porges, 1984; Porges, 1992a; Richards, 1987; Richards & Casey, 1991; Suess, Porges, & Plude, 1994) and verbal memory (Clark et al., 1999).

These findings establish a clear link between HRV measures and key aspects of affect and behavior regulation. The inhibition or self-control of immediate, automatic responses in order to achieve long-range goals is a key component of adaptive, flexible behavior. As elaborated in describing the TestEdge intervention in Chapter III, HRV feedback can play a significant role in helping students increase HRV and reduce both general and test-related anxiety. Through the use of such heart rhythm feedback, students can also learn to reduce incoherent heart rhythm patterns and correspondingly increase coherent patterns. This has multiple benefits, one of which is an increase in parasympathetic tone.

The Generation of Emotions: A Pattern-Matching Process

Further understanding of the heart's important role in emotional experience is gained by examining in greater depth how emotions are generated and processed in the brain and body. Recent years have seen the emergence of a new understanding of how the brain functions as well as of the brain-body dynamics involved in emotional processing. Rather than assembling thoughts and feelings from bits of data like a digital computer, the brain is an analog processor that relates whole concepts or *patterns* to one another and looks for similarities, differences, and relationships between them. This new way of understanding brain processes has also challenged long-held views of how emotions are generated. Psychologists once maintained that emotions were purely mental expressions generated by the brain alone. We now know, instead, that emotions have as much to do with the body as they do with the brain: thus, the emergence of emotional experience results from the ongoing interaction between the brain, the body, and the external environment.

Extensive work by eminent neuropsychologist Karl Pribram (Pribram, 1967, 1984, 1991; Pribram & Melges, 1969) has advanced the understanding of the psychophysiological processes that give rise to emotional perception and experience. Importantly, Pribram's theory of emotion elucidates a mechanism that accounts for the generation of emotions from the perception of body states, as described in Damasio's work, mentioned at the beginning of this chapter.

In Pribram's theory of emotion (Pribram & Melges, 1969), the brain is seen to function as a complex pattern-identification and matching system. According to his model, past experience builds within us a set of familiar patterns, which are instantiated in the neural architecture. Inputs to the brain from both the external and internal environments contribute to the maintenance of these patterns. Within the body, many processes provide constant rhythmic inputs with which the brain becomes familiar. These include the heart's rhythmic activity; digestive, respiratory and hormonal rhythms; and patterns of muscular tension, particularly facial expressions. These inputs are continuously monitored by the brain and help organize perception, feelings, and behavior.

Familiar input patterns form a stable backdrop, or reference pattern, against which new information or experiences are compared. When an input pattern is sufficiently different from the familiar reference pattern, a discontinuity or *"mismatch"* occurs. This mismatch, or *departure from the familiar pattern*, is what underlies the generation of feelings and emotions. In physiological terms, Pribram suggests that the low-frequency oscillations generated by the heart and bodily systems are the carriers of emotional information, and that the higher frequency oscillations found in the EEG reflect the integration, perception, and labeling of these body states along with perception of sensory input from the external environment. The mismatch between a familiar pattern and a pattern that is new or novel in either of these informational inputs is what activates emotional changes.

The background physiological patterns with which our brain and body grow familiar are created and reinforced through our experiences and the way we perceive and interact with the world. For example, a person living in an environment that continually triggers anxious or fearful feelings is likely to become familiar with those feelings, and with their neurological and hormonal correlates. In contrast, an individual whose experience is permeated by feelings of security, love, and care will become "familiar" with the psychophysiological patterns associated with those feelings. Importantly, once a reference pattern is established, in order to maintain stability, the neural systems attempt to maintain a match between the reference pattern, current inputs, and future behaviors. Since our psychophysiological systems are designed to maintain stability, returning to the familiar reference pattern gives us a sense and feeling of security, while remaining in unfamiliar territory causes unrest. Moreover , this is true even if the established reference pattern is one of chaos and confusion: *if the reference pattern becomes maladapted, the system will still strive to maintain a match to that pattern, even though it may be unhealthy or dysfunctional.*

This model thus provides a psychophysiological basis for understanding why patterns of chronic emotional stress and unhealthy behaviors can be so difficult to change. Through repeated experiences of stress, the brain learns to recognize the patterns of psychophysiological activity associated with "stress" as familiar, and therefore "comfortable." To the extent that these patterns of activity become part of our baseline reference, the system then automatically strives to maintain a match with these habitual psychophysiological and behavioral patterns, through a feedback process, despite their detrimental impact on health, emotional well-being and behavior. Thus, stress and self-defeating behaviors can become self-perpetuating and self-reinforcing without effective intervention.

However, as the system is in a dynamic relationship with its environment, this model also incorporates the means for change and development. Through a *feed-forward* process, like resetting a thermostat, as new input patterns are consistently experienced and thus reinforced in the neural architecture, they become familiar to the system, and the reference pattern is thus modified and fed-forward to a new set-point of stability. Once the new reference pattern is stabilized, the system then strives to maintain a match with inputs that characterize this new baseline.

Usually this process occurs automatically and unconsciously. *However, such a feed-forward, repatterning process can also be intentionally initiated*. This occurs as a pattern-matching operation in which the individual deliberately holds and projects a new emotional or behavioral pattern into the future as a "target of achievement," in Pribram's terms (Pribram, 1991). Holding the new pattern as a target in this way causes the psychophysiological systems to feed-forward as new patterns of input are experienced and processed. Essentially, the system makes continual adjustments in its patterns of activity until a match is achieved between the target and the current pattern of system activity. Eventually, if this process is sustained, a new baseline is created in which the new pattern is instantiated in the system as the reference pattern. *It is on this important principle that the HeartMath technology is based*. To further understand the processes by which these techniques work, it is necessary to examine the key role of the heart in this model.

More Than a Pump: The Heart's Key Role in the Emotional System

The model of emotion described here highlights the critical function of afferent input from the bodily organs to the brain in contributing to the input patterns that ultimately determine emotional experience. Although complex patterns of activity originating from many different bodily organs and systems are involved in this process, it has become clear that the heart plays a particularly important role. It is now understood that the heart is a key component of the emotional system, with the signals generated by its rhythmic activity playing a major part, moment by moment, in *determining the quality of our emotional experience* (McCraty, 2003; McCraty & Tomasino, 2006).

The heart is the primary and most consistent source of dynamic rhythmic patterns in the body. Furthermore, the afferent networks connecting the heart and cardiovascular system with the brain are far more extensive than the afferent systems associated with other major organs (Cameron, 2002). To add to this, it is now established that the heart is a sophisticated information encoding and processing center, with an intrinsic nervous system sufficiently sophisticated to qualify as "little brain" in its own right. Its circuitry enables it to learn, remember, and make functional decisions independent of the cranial brain, and its rhythmic input to the brain reflects these processes (Armour & Kember, 2004).

The heart also functions as a sensory organ, and is particularly sensitive and responsive to changes in a number of other psychophysiological systems. For example, heart rhythm patterns are continually and rapidly modulated by changes in the activity of either branch of the ANS, and the heart's extensive intrinsic network of sensory neurons also enables it to detect and respond to variations in hormonal rhythms and patterns (Armour & Kember, 2004). Finally, the heart is itself an endocrine gland that manufactures and secretes multiple hormones and neurotransmitters (Cantin & Genest, 1986; McCraty et al., 2006).

Thus, with each beat, the heart not only pumps blood, but also continually transmits dynamic patterns of neurological, hormonal, pressure, and electromagnetic information to the brain and throughout the body (McCraty et al., 2006). The multiple and continuous inputs from the heart and cardiovascular system to the brain, are, therefore, a major contributor in establishing the familiar reference pattern against which the current input of the moment of "now" is compared.

It also follows from this model that *changes* in the heart's patterns of activity can have an immediate and profound impact on emotional perception and experience. Such changes in heart activity most often occur unconsciously. However, one of the most important findings of our research is that changes in the heart's rhythmic patterns can also be *intentionally generated* (McCraty et al, 2006). This shift in the heart's rhythmic patterns is one of the physiological correlates of using the HeartMath coherence-building techniques (described in Chapter 3), which couple an intentional shift in attention to the physical area of the heart with the self-induction of a positive emotional state. We have found that this process rapidly initiates a distinct shift to increased coherence in the heart's rhythms. This, in turn, results in a change in the pattern of afferent cardiac signals sent to the brain, which serves to *reinforce* the self-generated positive emotional shift, making it easier to sustain. Often this shift is also associated with enhancements in perception and cognition that enable more effective reasoning, decision making, and action when confronted with stressful or challenging situations (McCraty & Tomasino, 2006).

While such a positive emotion-driven shift in heart rhythm activity clearly leads to immediate benefits by helping to transform stress in the moment it is experienced, it can also contribute to long-term improvements in emotion regulation abilities and emotional well-being that ultimately affect many aspects of one's life. This is because each time an individual intentionally self-generates a state of heart rhythm coherence, the "new" coherent pattern—and "new" repertoire for responding to challenge—are reinforced in the neural architecture. With consistency of practice, these patterns become increasingly familiar to the brain. Thus, through the feed-forward process Pribram describes, these new, healthy patterns become established as a new baseline or reference, which the system then strives to maintain. It is in this way that HeartMath tools facilitate a *repatterning process* whereby the maladaptive patterns that underlie the experience of stress are progressively replaced by healthier physiological, emotional, cognitive, and behavioral patterns as the "automatic" or familiar way of being (McCraty & Tomasino, 2006).

Relational Context of Anxiety and Poor Performance in School

Before moving on to describe the tools and techniques employed in the TestEdge intervention, we show how the nature and organization of the socioemotional relations at school and in the classroom directly affect the quality and intensity of emotions that students experience. Because of this important two-way link between relations and emotions, how the socioemotional environment is organized can enhance or inhibit student learning and psychosocial growth.

At a sociological level, the nature and quality of socioemotional relations in the classroom and in the school not only strongly influence the well-being and performance of the individual but also shape the function and effectiveness of the school as a whole. It is not surprising, therefore, that recent research indicates that high incidences of students with aggressive and other behavioral problems, high student dropout rates, high teacher burnout and attrition rates, and low academic performance are all produced in schools which, as social systems, are unable to provide a positive, nurturing environment of socioemotional support and guidance for students and teachers (e.g., Cohen, 2006; Greenberg et al., 2003; Sanders & Horn, 1998).

A Relational Approach

The basic proposition is that all functionally healthy social collectives—social relationships, families, schools, communities, organizations, even societies—are comprised of two systems of socioemotional relations that interconnect all individuals in a collaborative order of mutually beneficial bonds (see Bradley, 1987 & 2004; Bradley & Pribram, 1998). The first system comprises the resonance relations of positive affective attachment. These relations are the emotional means of connection and communication among people and throughout the social collective as a whole. The second system comprises the relations of social regulation and guidance. These relations of social control provide the means to direct, coordinate, and integrate the actions of all subunits and individuals within the collective toward the achievement of a mutually beneficial collective goal. In short, such an order of cooperative attachment and social regulation is a prerequisite to efficient bio-emotional and resource utilization, and effective communication and to individual and collective accomplishment and development (Bradley, 2004).

When relations of positive emotional attachment and social regulation are reciprocally coupled in an adaptive order of constantly coevolving cooperative organization, a dynamic stable order of social coherence emerges in the system as a whole, with resulting optimal communication, psychosocial development, and effective action (Bradley, 2004). At the individual level, as Piaget has shown (1965/1995), the "logic of cooperative interaction" is the optimal social form for learning, intellectual growth, and social development: it not only creates the capacity for reason, logic, and the ability to discern objective truth, but it also generates the requisite psychosocial skills and experience for effective and adaptive collaboration.

Conversely, when there is insufficient connection among all individuals on these two systems of relations, or when the two systems of relations are not structured as a reciprocal coupling of cooperative interaction, an incoherent social structure emerges. This impedes the free flow of communication to produce less than optimal or even dysfunctional outcomes. For the individual, social incoherence—relations involving prolonged tension, conflict, or disaffection—is ineffective as a social means to facilitate learning and growth. This increases stress and anxiety, particularly when interactions are infused with negative emotional energy. Under these conditions, individual development is blunted and can even regress or be pathologically stifled if such interaction continues over the long term (Bradley, 2001).

The emotional states of individuals and the socioemotional field of the collective—such as the school and the classroom—are interrelated and act to amplify and mutually reinforce each other (Bradley, McCraty, & Rees, 2004). For example, as individuals increase psychophysiological coherence, through the activation of sincere positive emotions such as appreciation and care, socioemotional attunement may be increased. This, in turn, increases the coherence and harmony of the relations and interactions within the group. Similarly, the creation of a coherent field of social relations by a group may facilitate the generation and maintenance of positive emotions and psychophysiological coherence in its members. These processes of mutual reinforcement produce stable, effective collective function, which enhances psychosocial wellbeing and individual growth and performance.

Conversely, discordant emotions and relations also act to strengthen each other as a mutually reinforcing cycle. When individuals in a group are in an incoherent psychophysiological state, activated by stress and negative emotions, this discordance permeates the interactions and relations among them. This produces an incoherent order of relations within the group. By the same token, the creation of an incoherent field of relations in the group will make it difficult for the individual members (even more so children and adolescents) to self-activate and sustain a coherent emotional and psychophysiological internal state. The amplification of incoherence at both the psychophysiological level of the individual and the relational level of the group increases stress and anxiety and reduces task performance and group effectiveness (McCraty, Bradley, & Tomasino, 2004–2005).

Socioemotional Dynamics of the School

As holds for the growth and development of infants and preschoolers (Schore, 1994; Hinde, 1992), learning requires that students be placed in a nurturing environment of positively charged emotional bonds in which they feel sufficiently safe and secure to explore themselves, take risks, and be open to adult inspiration and guidance to inform their growth and development. This is possible so long as the students remain fully engaged in the fabric of socioemotional bonds that connect them to the school.

In dysfunctional school systems where the social environment is fraught with tension, disaffection and even hostility and conflict, student stress and anxiety are a direct consequence of the breakdown of relations of emotional attachment and relations of leadership and guidance. In such circumstances, the students' behaviors reflect an increasing level of disaffection with and disengagement from the school as a social system. Thus, tardiness—being late for school or class—and minor disciplinary infractions (disrupting class) are indicative of an initial lack or loss of positive socioemotional connection, whereas chronic truancy (skipping class or school) and serious disciplinary infractions (abusive and/or violent behavior) reflect an alienated state which manifests in extremely negative feelings and behaviors such as hostility, anger, and rage.

The degree to which students are able to learn is strongly influenced by the quality of the socioemotional bonds that connect them to their teachers and their classmates, and to the school community as a whole. Under optimal conditions, positive relations of caring and attachment, in combination with enlightened and caring firm leadership and guidance, are reciprocally coupled into an order of mutually beneficial cooperative relations so that the dynamic of this relationship facilitates and reinforces student learning. However, in dysfunctional schools, where the relations of positive attachment are weak and are replaced by negative interactions of tension, disaffection, and even conflict, and the relations of guidance are directed solely to discipline and control, increased stress and anxiety occur and poor learning and academic performance are highly likely results.

Socioemotional Dynamics of the Optimal Classroom

The difference in the relational organization and consequences of the dynamics of these two situations can be illustrated in the two classroom settings depicted in Figure II.5. As shown in Figure II.5a, the relational dynamics based primarily on bonds of positive emotions create a field of socially coherent relations that is organized for optimal communication. These effects extend beyond the teacher–student dyad and produce positive socioemotional changes throughout the network of relationships among students in the classroom, as depicted in the model in Figure II.6. While not shown in the figure, it is also likely that these field effects will radiate outwards, thereby also enhancing the quality of relations in the school, among friends and family, and in the community (Bradley, McCraty, & Rees, 2004).

By contrast, the pattern for the second classroom, shown in Figure II.5b, is generated by emotional tension, stress, and conflict originating from both within and outside the classroom. This leads to discordant relations charged primarily with negative emotions. The resulting interactional dynamics produce an incoherent socioemotional field. In this less than optimal situation, effective communication is impeded, and a negative impact on student learning and growth is likely.

Figure II.5. Model Showing the Difference in Organization of Socioemotional Energy Fields in Coherent Interactions and Incoherent Interactions among Teacher and Students

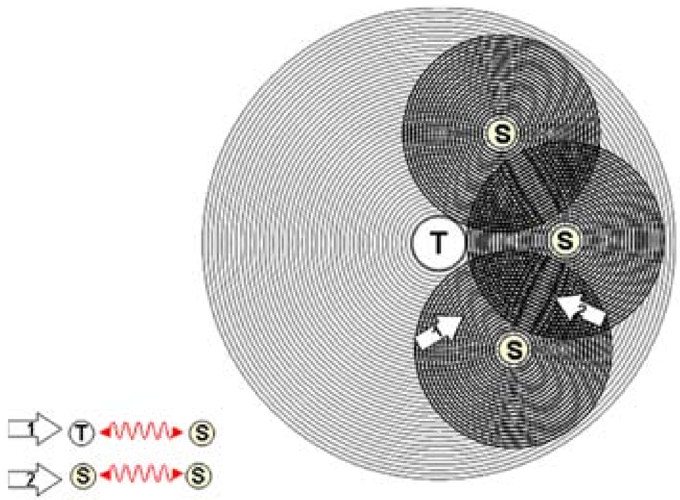


Figure II.5a: Coherent Interactional Dynamics

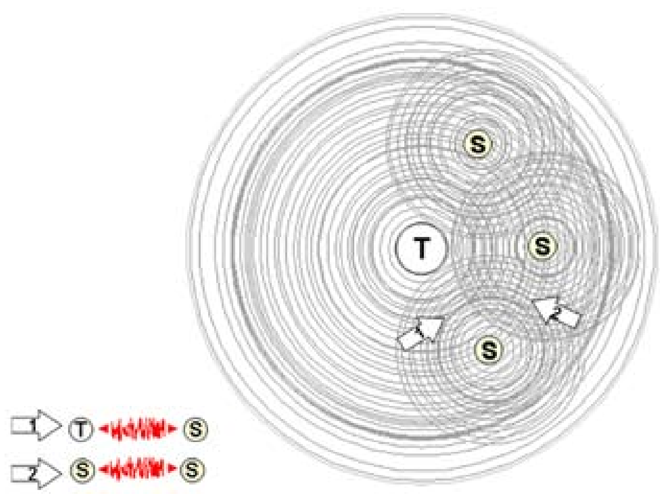


Figure II.5b: Incoherent Interactional Dynamics

Figure II.5. Each set of concentric circles represents the bioemotional energy field generated by the respective teacher (T) or student (S). The circles are waves of bioemotional energy that radiate outwards from the individual in the same way that a pebble generates a succession of waves when dropped in a pool of water. Note that in *coherent interactions* (Figure 5a), the distance between the wavefronts (the circles) in all four sets of concentric circles is identical, indicating a uniform frequency and pattern for each set of waves. Because the interaction between the radiating waves is synchronized and harmonious (in phase), the resulting interference pattern creates a *coherent channel of connection*, which is optimal for communication between the teacher and the students (see the area indicated by the arrow labeled # 1) and between students (arrow labeled # 2). By contrast, in *incoherent interactions* (Figure 5b) the distance between the wavefronts in all sets of concentric circles is highly variable. Because the interaction between the radiating wavefronts is unsynchronized (out of phase), the resulting interference pattern creates an *incoherent channel of connection*, which impedes effective communication and information transmission.

Figure II.6. Model of a Classroom with Coherent Socioemotional Energy Fields Showing Emergent Channels of Optimal Communication Among Members of the Class

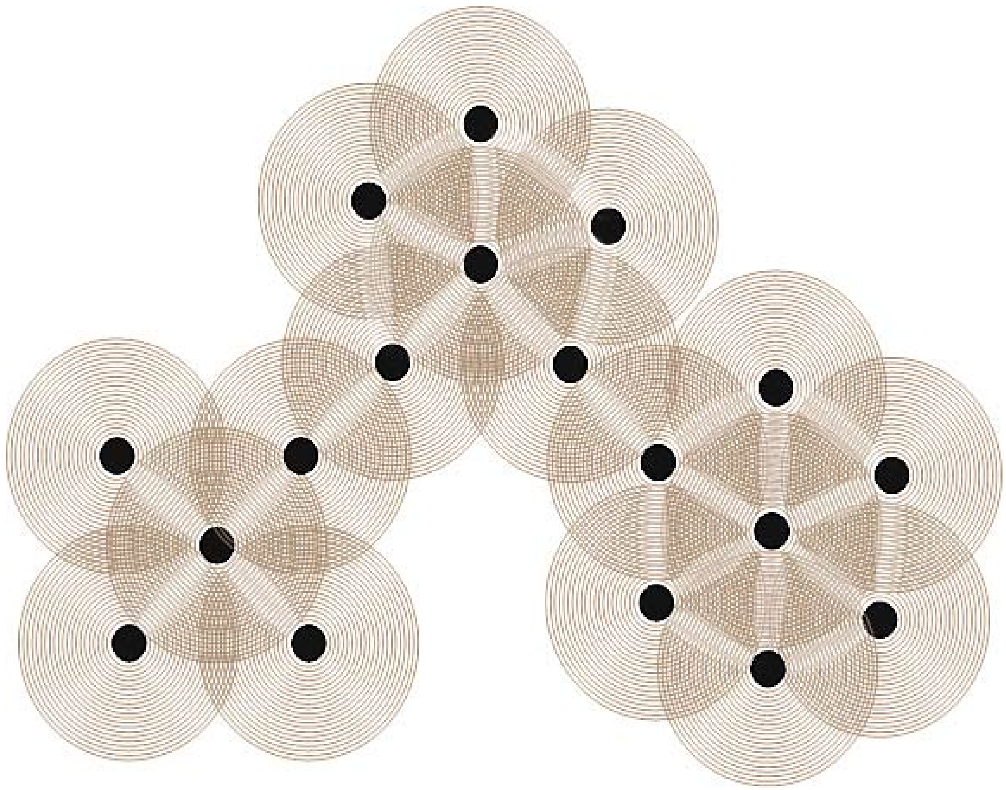


Figure II.6. When all individuals in a classroom are socioemotionally attuned to one another at the same frequency of bioemotional energy, a wave field of *coherent interactions* is generated, in which waves of energy (concentric circles) radiating from each individual are synchronized and harmonious (in phase). The resulting interference patterns create an emergent network of *coherent channels of connection* interlinking all individuals, which is optimal for communication and relational connection.

Based on the relational dynamics depicted in this model, an effective teacher is an agent who facilitates a sustained shift to high levels of emotional and social coherence in the field of relations among classroom members (Bradley, McCraty, & Rees, 2004). By activating and aligning high levels of positive emotional energy, the teacher facilitates the creation of a field of coherent socioemotional relations among the members of the classroom. These coherent connections create a channel for optimal communication, which in turn enables a qualitative shift in student intellectual, emotional, and social development. It is expected that if sustained and socially reinforced, the effects of this relationship will have positive consequences for student growth, as reflected in measures of academic achievement and psychosocial development.

Creating a healthy social and emotional environment that optimizes both learning and growth ultimately requires a school-wide approach in which positive emotional connections are infused with caring, firm leadership to create a system-wide fabric of cooperative relations. The more attached and attuned students, teachers, and staff are to one another, the greater the social coherence of relations in the system as a whole, and the greater the reduction of stress and anxiety. Under these conditions, learning, development, and performance for all individuals are likely optimized. In short, building and nurturing the bonds of socioemotional attachment and relations of guidance that connect individuals to one another and to the school system are essential to creating schools that produce students prepared to take their role in society as creative and productive citizens (Cohen, 2006; Greenberg et al., 2003).

Chapter III
The TestEdge Intervention: An Emotion-Based Approach

In the previous chapter we saw that emotions often play a fundamental and even independent role in activating the stress response. However, despite the fact that most stress has an emotional source, it is interesting to observe that *most of the widely used stress management interventions do not directly focus on emotions*.

Common Approaches to Stress Management and Anxiety Reduction

Relaxation has long been seen as the ultimate remedy for stress and anxiety (Benson, 1975). Although relaxation techniques can be helpful and beneficial in that they temporarily draw attention away from distressing feelings and thereby reduce physiological arousal, it is often assumed that by merely reducing the negative physiological reaction to stress and anxiety, an automatic positive emotional shift will occur. However, recent research challenges this assumption. For example, there is evidence that the positive and negative aspects of experience are mediated by separate psychological systems rather than being opposite ends of a single dimension (MacLeod & Moore, 2000). Furthermore, arousal-reducing techniques such as relaxation generally do little to transform an individual's underlying feelings about the source of stress; therefore, the stressful feelings are likely to recur.

Other techniques commonly used to manage stress are derived from cognitive-behavioral psychotherapy (Beck, 1976; Greenberger & Padesky, 1995). These interventions are based on the cognitive perspective described earlier—that negative thoughts drive unhealthy reactions and behaviors, and that these thoughts should therefore be the focus of therapeutic intervention. In accord with this model, although emotions may be explored, they are seen as a consequence of maladaptive thoughts. Cognitive refocusing or reframing techniques are often effective in facilitating a cognitive or conceptual shift, which can be important. However, the understanding of the interactions between the emotional and cognitive systems presented above may help explain why such techniques are often less than fully effective in modifying the emotional patterns at the root of much anxiety and stress.

Most cognitively-based programs for reducing test anxiety and improving test performance focus mainly on test-taking strategies—cognitive techniques for improving memory and recall, comprehension, and reasoning and for maintaining intellectual focus and mental self control (Beidel, Turner, & Taylor-Ferreira, 1999; Roth, Paris, & Turner, 2000). Typically, however, they do not address the *feelings* stemming from the emotional component of test anxiety and the psychophysiological repercussions of such feelings. As a result, cognitive-focused test-taking strategies alone may do little to lessen the emotional feelings of fear and anxiety and the associated autonomic and neuroendocrine responses.

Overview of the TestEdge Program

The new research we reviewed in the previous chapter, showing how the cognitive and emotional systems interact, has significant implications for test anxiety interventions: it suggests that *intervening at the level of the emotional system itself* is a more direct, efficient, and effective way to modify the maladaptive emotional patterns underlying the physiological, psychological, and behavioral manifestations of the anxiety response.

The TestEdge program used in this study is an emotion-focused intervention, employing a set of techniques that utilize the *activation of positive emotions* to help shift test anxiety and other sources of emotional stress (described in further detail below). A growing body of research is beginning to provide objective evidence of the centrality of positive emotions to optimal functioning in nearly all spheres of human experience. Positive emotions and their appropriate regulation have been shown to be critical to neurobiological development in infants and children (Schore, 1994). Positive emotions have been shown to affect the way we think and address challenges—enhancing many aspects of cognition and performance, including cognitive flexibility, creativity, receptivity, and innovative problem solving (Fredrickson, 2002; Isen, 1999; Isen, Daubman, & Nowicki, 1987). Positive emotions further shape our behavior, promoting helpfulness, generosity, and effective cooperation (Isen, 1987). When linked as a balanced coupling with relations of social control, relations of positive affect have been shown to be the basis of group stability and effective collective function (Bradley, 1987 & 2001; Bradley & Pribram, 1998). In short, research suggests that actively cultivating positive emotions is critical to people's emotion regulation, effective adaptation to challenges, individual growth and development, and collective social function (Aspinwall, 1998; Bradley, 2004; Fredrickson, 2002; McCraty & Childre, 2004).

The HeartMath system of emotion regulation tools and techniques taught in the TestEdge program is based on a large body of research which demonstrates that *sustained positive emotions facilitate an emergent global shift in psychophysiological functioning,* which is reflected in a distinct change in the pattern of heart rhythm activity (McCraty et al., 2006). This global shift generates a psychophysiological state that is optimal for learning and test-taking. This state, termed *psychophysiological coherence*, is characterized by increased synchronization, harmony, and efficiency in the interactions within and among the physiological, cognitive, and emotional systems. Physiologically, the coherence state is marked by the development of a smooth, sine-wave-like pattern in the heart rate variability trace (*heart rhythm coherence*; see Figures II.3 and III.1), which reflects increased order in higher-level control systems in the brain, increased synchronization between the two branches of the ANS, and a general shift in autonomic balance towards increased parasympathetic activity. Studies have shown that self-induction of the coherence state is associated with reduced stress and anxiety, increased emotional stability, and significant improvements in cognitive performance on tasks requiring focused attention, discrimination, and long-term memory (McCraty et al., 2006).

Based on this research, the Institute of HeartMath developed a set of positive emotion refocusing and emotional restructuring techniques that help people learn to self-generate and sustain psychophysiological coherence and its associated benefits (Childre & Martin, 1999; Childre & Rozman, 2005, 2006). In brief, HeartMath techniques combine a shift in the focus of attention to the area around the heart (where many people subjectively feel positive emotions) with the intentional self-induction of a sincere positive emotional state, such as appreciation. We have found that such a shift in focus and feeling experience facilitates the natural emergence of the psychophysiological coherence state. This shift to coherence increases nervous system synchronization and results in a change in the pattern of afferent cardiac signals sent to the cognitive and emotional centers in the brain. In turn, this serves to interrupt or prevent the triggering of the body's normal stress response and facilitates higher cognitive faculties and emotion regulation abilities that are normally compromised during stress or negative emotional states. This sharpens one's discernment abilities, increasing resourcefulness and often enabling problematic issues, interactions, or decisions to be assessed and dealt with from a broader, more emotionally balanced perspective (McCraty & Tomasino, 2006).

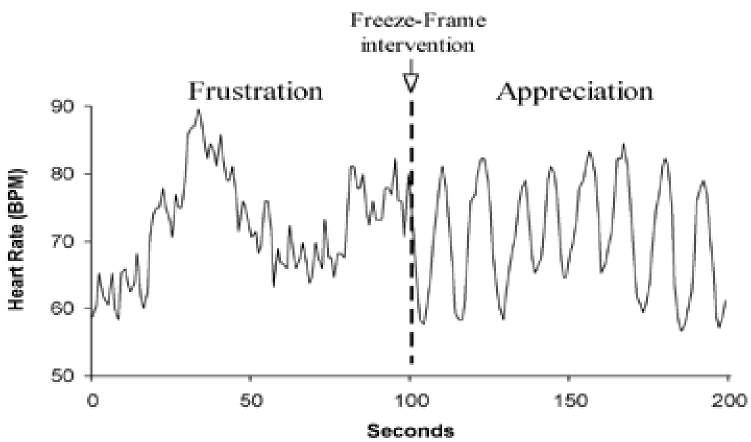


Figure III.1. Shift to Coherence.
The real-time heart rate variability (heart rhythm) pattern is shown for an individual making an intentional shift from a self-induced state of frustration to a genuine feeling of appreciation by using the Freeze-Frame positive emotion refocusing technique (at the dotted line) developed by the Institute of HeartMath. Note the immediate shift from an erratic, disordered (incoherent) heart rhythm pattern associated with frustration and emotional stress to a smooth, harmonious, sine-wave-like (coherent) pattern as the individual uses the positive emotion refocusing technique to self-generate a feeling of appreciation.

To understand what happens at the emotional level, it is necessary to draw on Pribram's model of emotion (Pribram & Melges, 1969), which was described in detail in Chapter II. As elucidated by Pribram, the movement to a more organized pattern of cardiac afferent input that accompanies a coherent heart rhythm pattern is one that the brain associates with feelings of security and well-being, resulting in a "pattern match" with positive emotional experience. This shift in the pattern of the heart's input to the brain thus serves to *reinforce* the self-generated positive emotional shift, making it easier to sustain. Through consistent use of the HeartMath tools, the coupling between the psychophysiological coherence state and positive emotion is further reinforced. This subsequently strengthens the ability of a positive feeling shift to initiate a beneficial physiological shift towards increased coherence, or a physiological shift to facilitate the experience of a positive emotion (McCraty & Tomasino, 2006).

While positive emotion-based coherence-building techniques have been shown to be effective in helping individuals reduce emotional stress *in the moment* it is experienced, such techniques have also been demonstrated to lead to enduring improvements in emotion regulation abilities that affect many aspects of one's life. As expected by Pribram's model of emotion, our research indicates that the intentional application of these coherence-building techniques, on a consistent basis, effects a *repatterning process* whereby coherence becomes increasingly familiar to the brain and nervous system. In this way, the state of coherence becomes progressively instantiated in the neural architecture as a new, stable psychophysiological baseline or set-point (McCraty & Tomasino, 2006). Thus, through this feed-forward process, once the coherence state is established as the familiar pattern, the system then strives to maintain this new set-point or stable state automatically. The result is that unhealthy or maladaptive psychophysiological patterns that underlie the experience of stress are progressively replaced with ones that foster increased physiological efficiency, mental acuity, and emotional stability. This repatterning process thus renders coherence a more readily accessible state during day-to-day activities, even in the midst of stressful or challenging situations such as test-taking.

Studies conducted across diverse populations in laboratory, organizational, clinical, and educational settings have demonstrated that HeartMath positive emotion-focused techniques are effective in producing both immediate and sustained reductions in stress and anxiety, together with improvements in many dimensions of psychosocial well-being (Arguelles et al., 2003; Barrios-Choplin, McCraty, & Cryer, 1997; Luskin et al., 2002; McCraty et al., 1995; McCraty, Atkinson, & Tomasino, 2003; McCraty et al., 1999; McCraty et al., 1998). Moreover, these interventions have also been shown to give rise to significant improvements in key health and performance-related measures (for summaries, see McCraty, Atkinson, & Tomasino, 2001; McCraty & Childre, 2004; McCraty & Tomasino, 2006). Collectively, results indicate that such techniques are easily learned and used, produce rapid improvements, have a high rate of compliance, and are readily adaptable to a wide range of sociocultural and demographic groups.

Previous studies of the use of HeartMath interventions in educational settings have demonstrated significant reductions in measures of psychosocial stress and improvements in attitudes, behaviors, and test performance. For example, the Miami Heart Research Institute examined the impact of a HeartMath program on psychosocial functioning and physiological responses to stress in students at a Florida middle school (McCraty et al., 1999). Results indicated significant improvement in 17 of the 19 areas of psychosocial functioning examined using the Achievement Inventory Measure (AIM). AIM measures include three areas of emphasis: Achievement Aptitude, Interpersonal Skills, and Mental Attitudes. Moreover, the study showed that students were able to use the self-regulation skills to favorably modulate their physiological response to acute emotional stress in real time, thus demonstrating increased physiological stress resilience in relation to a control group. In another study conducted in the Minneapolis Public School System, high school seniors who received a three-week training in HeartMath tools demonstrated substantial improvements in test scores and passing rates on Minnesota's state-required exit examinations in math and reading. Students also experienced significant reductions in self-reported assessments of hostility, depression, and other measurements of psychosocial stress (McCraty et al., 2000).

The TestEdge Intervention

The overall intervention consisted of three primary components involving both teachers and students in the intervention school:

- **The Resilient Educator program.** A professional development program designed for teachers, providing HeartMath coherence-building tools and strategies for reducing stress and burnout and increasing effectiveness in the classroom.
- **The TestEdge program.** A learning program providing HeartMath coherence-building tools, concepts, and strategies to help students reduce test anxiety while enhancing their emotional self-regulation skills and psychosocial functioning.
- **Heart rhythm coherence feedback (Freeze-Framer system).** A computer-based interactive system designed to facilitate practice of the coherence-building techniques taught in the TestEdge and Resilient Educator programs by providing

real time feedback of heart rate variability (heart rhythm) patterns and quantifying the level of heart rhythm coherence achieved.

Resilient Educator

The Resilient Educator program was delivered to teachers participating in the study at both the primary and secondary sites to provide them with an understanding and personal experience of the HeartMath tools and techniques prior to their beginning classroom instruction of the TestEdge program.

At the intervention school, the 10th grade English Language Arts teachers attended a one-day Resilient Educator program in October of 2004, several months before they were to begin teaching the TestEdge program. This was done to give the teachers ample time to practice and internalize the tools and techniques in their own lives prior to teaching them to students. The Resilient Educator program provides instruction in several HeartMath coherence-building techniques to help teachers reduce stress, burnout, and negative emotions; rekindle energy and motivation; improve communication; and increase classroom coherence and effectiveness. The workshop includes the Neutral, Freeze-Frame, Heart Lock-In, and Effective Communication techniques (Childre & Martin, 1999; Childre & Rozman, 2006); a series of exercises designed to help participants understand how to apply these techniques in the classroom and in day-to-day life; a science module explaining how emotions and heart–brain interactions are directly involved in the psychophysiology of learning and performance; and instruction in the use of the Freeze-Framer Interactive Learning System (now emWave PC), a computer-based heart rhythm monitor and coherence-building training system. In addition, each teacher in the program was given a Freeze-Framer system for personal use at home and in the classroom.

In January of 2005, in preparation for teaching the student program, the teachers attended a second one-day training workshop in order to help familiarize them with the specific concepts and techniques taught in the TestEdge high school program. The teachers were provided with a teacher's manual, video, CDs, posters, and visual aids for helping to introduce the TestEdge program in the classroom. They were also given an opportunity during the workshop to practice with the Freeze-Framer system.

TestEdge

The TestEdge program (Institute of HeartMath, 2004) teaches students how to intentionally self-regulate their emotional and physiological responses to challenging and stressful situations through the use of positive emotion-based tools and techniques that facilitate the activation of the psychophysiological coherence state. The program teaches students how to apply HeartMath coherence-building tools and technologies in test preparation and test-taking; to facilitate learning and increase retention of academic material; to increase emotional self-awareness; and to more effectively handle stress and challenges, both at school and in their personal lives.

The program includes the following materials: a Teacher's Manual, a CD containing PowerPoint and audiovisual presentations used to illustrate key concepts, and a Student Workbook for each student. In developing the 63-page Student Workbook, a specific effort was made to present the material in an engaging way and in sufficient depth so as to enable students to learn and benefit from the program irrespective of the teacher's instructional ability or level of commitment in teaching the program, given time constraints or other factors.

The content of the TestEdge program is organized in 12 lessons, with each lesson requiring around 20 minutes. The program includes the Neutral, Attitude Breathing, and Freeze-Frame techniques (Childre & Martin, 1999; Childre & Rozman, 2006) as well as lessons on goal setting, the basic physiology of the three major divisions of the brain and their functions, emotional memories, the physiology of positive and negative emotions, basic knowledge of heart–brain communication and how this communication affects performance, common performance blocks, and how to apply the HeartMath techniques in various aspects of the learning process, including test preparation and test-taking. As part of the program, students were also taught how to use the Freeze-Framer coherence-building system (described further, below) and given the opportunity to practice with this technology both in and outside of class.

In the experimental school, the TestEdge program was delivered by the English teachers during the normal class period starting in January 2005 and ending in May 2005. Typically two lessons were taught each week. To help facilitate a smooth integration of the program, Institute of HeartMath staff members co-taught the first lesson of the TestEdge curriculum and the first Freeze-Framer lab session with the teachers.

The TestEdge program was also introduced in the secondary study schools, with elementary, middle, and high school versions appropriate for each level.

Heart Rhythm Coherence Feedback – Freeze-Framer Interactive Learning System

Both the teacher and student programs described above incorporated training with the Freeze-Framer Interactive Learning System[2] (Quantum Intech, Inc., Boulder Creek, CA), a unique heart rhythm coherence feedback system, designed to facilitate learning and implementation of the self-regulation and anxiety reduction techniques (McCraty, 2005). Through noninvasive measurement of the pulse, this personal computer-based system displays a user's changing heart rhythm patterns (heart rate variability) in real time and quantifies the level of heart rhythm coherence, the key marker of the psycho-physiological coherence state. This technology enables individuals to see and feel for themselves how anxiety and different emotions affect their physiology and facilitates learning the positive emotional shifts associated with coherence. As users practice the coherence-building techniques, they can readily see and experience the changes in their heart rhythm patterns, which generally become smoother and more sine-wave-like as users shift from negative or stressful feelings to a positive emotional state (see the real-time example shown in Figure III.1). The Freeze-Framer system also includes a comprehensive audiovisual tutorial in the HeartMath coherence-building techniques, three interactive games whose outcomes are determined by the level of coherence achieved, and a multi-user database to store results and track progress over time. This system has been used effectively in many educational settings, with students of widely diverse sociocultural backgrounds and academic levels, and has been found to be a fun and engaging way to reinforce student emotion regulation skills.

The Freeze-Framer system was installed in the experimental school's three computer labs where students had two of their TestEdge lessons in order to practice the tools they had learned in the classroom. One of the computer labs was located in the school library; this afforded students additional opportunity to practice with the Freeze-Framer system before or after school.

As with other aspects of the program, the Freeze-Framer system was also introduced in the secondary school sites.

[2]The Freeze-Framer system has since been updated and renamed as "emWave® PC."

Chapter IV
Research Design and Methodology

The major proposition that informs this study is that the ability of individuals to self-regulate their emotional responses in high-stakes testing situations is related to the degree of performance anxiety they experience when taking important tests. In turn, the level of their test anxiety has a direct and independent effect (all other things being equal) on their ability to perform at their true level of competence on tests. The psychophysiological basis of this proposition is that the higher the intensity of negative emotions, the greater the activation of the autonomic nervous system and the greater the degree of disorder occurring in higher brain function and nervous systems dynamics. The consequences of this psychoneurophysiological process results in impaired perception, cognition, intellectual function and intentional behavior (Bush, 1989; Evers-Lush, 1991; Fanelli, 1991)

Study Hypotheses

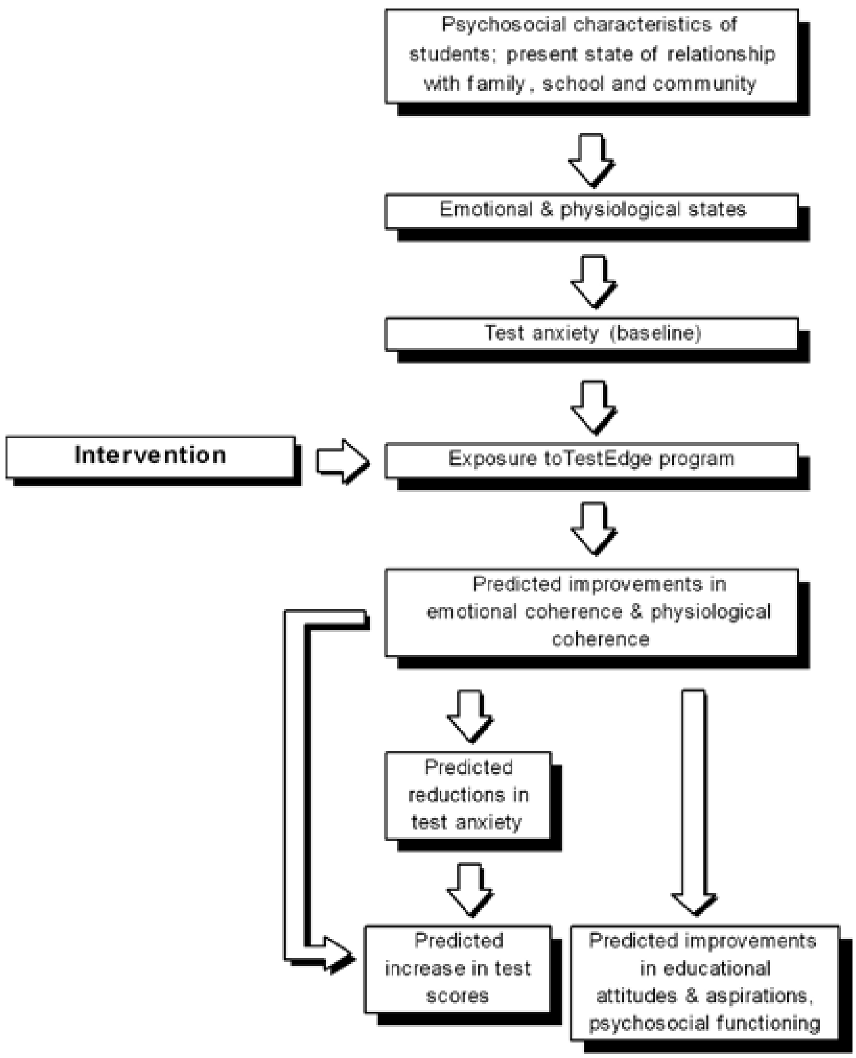
Two major hypotheses were tested in the study. We expected that—all other things being equal—

> **Hypothesis 1:** Competence in practicing the TestEdge tools would result in significant improvements in emotional self-regulation and psychophysiological coherence. These changes would produce a marked reduction in test anxiety, which in turn would lead to a corresponding improvement in academic and test performance.

Hypothesis 2: Given the logic of Hypothesis 1, we also expected that there would be concomitant improvements in emotional well-being and stress management, life aspirations, behaviors, and relationships in the student's life as a whole, as well as improvements in classroom climate, organization and function.

The causal logic is summarized in Figure IV.1.

Figure IV.1. Expected Effects of the HeartMath Intervention on Student Psychophysiological States, Test Anxiety, Attitudes, Social Behavior, and Test Performance



Primary Study Research Design

The primary study focused on an in-depth investigation of students at the 10th grade level in two high schools with the TestEdge intervention being introduced to one school and the second school serving as a control. For robust measurement, we employed a multi-methods approach utilizing standardized measurement instrumentation and data collection supplemented with qualitative observation, as described below. The Institutional Review Board (IRB) approval for this project was obtained through Claremont Graduate University, Claremont, California.

Study Design

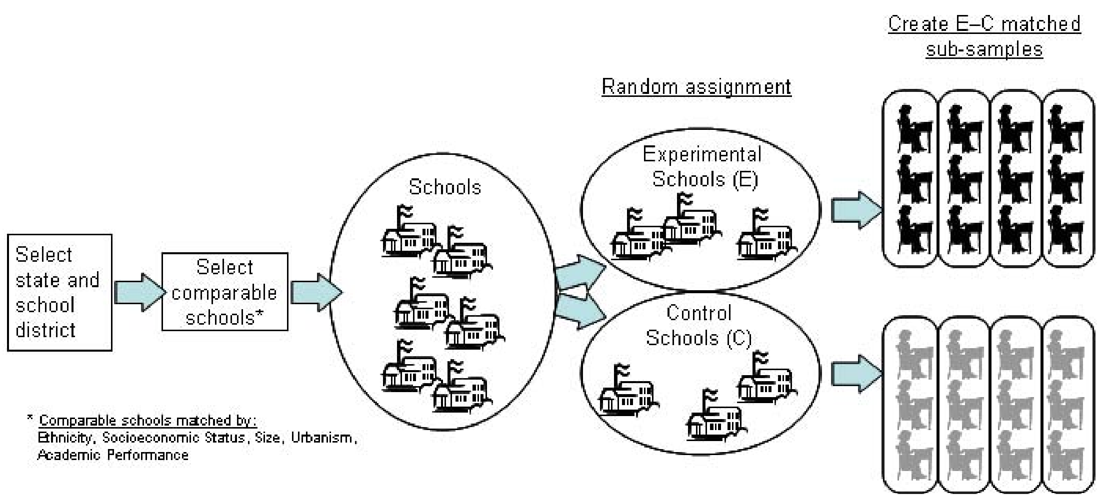
The study utilized a randomized, quasi-experimental, longitudinal research design, as shown in Figure IV.2. This design involves matching natural social groups (schools) on certain significant characteristics to achieve some degree of control in an otherwise "open" field setting, introducing an intervention in one group, and gathering pre- and post-intervention measurements.

Figure IV.2. Logic of Experimental Design

	Pre	Intervention	Post
Experimental group: R	O_1	X^{HM}	O_2
Control group: R	O_3	--	O_4
R = random assignment; O = observation or measurement; X^{HM} = intervention with TestEdge			

Because the research is being conducted in an open field setting—viz., 10th grade high school classrooms—there is not sufficient control to meet the required conditions in a controlled experiment for strict equivalence between experimental and control groups achieved by rigorous random assignment of individual cases into each group from a common population. Instead, control on key variables is achieved by using a matched-groups approach in selecting the schools, and then randomly assigning the sample of matched schools to experimental and wait-list control groups. Within each school it was decided to study the whole population of 10th grade students, for reasons given in the "Sample Selection" section that follows. For the physiology study, a random sample of 136 students—stratified on Test Anxiety, Gender, and Class Academic Level, and divided between the experimental and control schools—was selected from the usable sample of 10th grade students. For certain analyses of sub-populations within the schools, matched groups of students on certain sociodemographic characteristics and matched groups of classrooms were also constructed, as depicted in Figure IV.3. By keeping each school geographically separate as either an experimental or a control site, the spillover effects of the intervention in experimental classrooms on other students in the school via friendship networks can be minimized. This way a change in school culture due to the effects of the intervention was confined to the experimental school and was unlikely to spill over and affect students in the control school.

Figure IV.3. Logic of Sampling Process



Sample Selection

The 10th grade level was selected over the middle and elementary school levels, due to the greater emotional stability and maturity of students, the increased significance of high-stakes tests for this population, the availability of age-appropriate measurement instrumentation, and the national importance of improving high school student test performance on graduate exit examinations and achievement tests. In California, the 10th grade is the first year in which students are required to take the California High School Exit Examination (CAHSEE).

Two high schools in Northern California were selected and agreed to participate in the study: One school was located in the North Central Coast (hereafter referred to as "Valley High School" to preserve confidentiality) and the other was located in the San Francisco Bay area (hereafter referred to as "Bay High School"). Using the latest publicly available data (2003–2004) at the time of sample selection, "Valley High School" had a total enrollment of 2,840 students; we assumed that approximately twenty-five percent (~710 students) were in the 10^{th} grade. In terms of ethnic composition, just over half (53.6%) were Hispanic or Latino and 42.3% were "White not Hispanic." The remaining students were Filipino (1.6%), Asian (1.2%), "African American not Hispanic" (0.8%), and American Indian or Alaska Native (0.4%). Students classified as "English Learners" constituted 5.3% of the student population. Almost one-in-five students (18.4%) qualified for free or reduced-price meals. The school's Academic Performance Index (API) was 672.

Bay High School's 2003–2004 total enrollment was 1,928 students, of whom approximately 482 students were in the 10^{th} grade, given our assumption of 25%, above. In terms of ethnic composition, more than half (55.9%) were "White not Hispanic," while approximately one-fifth (22.1%) were Hispanic or Latino, and one-tenth (10.7%) were Asian. The remaining students were "African American not Hispanic" (4.5%), Pacific Islander (3.3%), Filipino (1.3%), and American Indian or Alaska Native (0.5%). Students classified as "English Learners" constituted 14.1% of the student population. One-tenth (10.2%) of the student population qualified for free or reduced-price meals. The school's Academic Performance Index (API) was 709.

Both schools' students took the CAHSEE in March and the California Standards Test (CST) in April. The school sites were chosen based on their relatively similar profiles, convenience of location to the Institute of HeartMath, willingness to participate, and adequate sample size for the study, which was a significant limiting factor.

The random selection of the intervention and control schools was determined using a computer program. "Valley High School" was selected as the experimental site and "Bay High School" as the wait-list control site. The study was conducted over the 2004–2005 academic year, with the students at the control school receiving the TestEdge program in the 2005–2006 academic year. All quantitative and qualitative measurements were conducted in both the experimental and control schools during the same time periods. Parental and student consent forms were obtained for all students participating in the study. Study participants were assigned a randomly selected personal identification number that was used for all data collection, coding, recording, and analysis of the database.

Epistemological Considerations[3]

From an epistemological perspective, the TENDS research team believed that it was important to investigate both the objective and subjective dimensions of the research topic and study participants. Thus, it was decided to complement the more quantitative research approach of the IHM team with the more qualitative interpretative inquiry approach of the CGU team. Adhering to the more traditionally defined principles of objectivity, reliability, and validity, the IHM team aimed to use formal measurement procedures (questionnaires, formal interviews, standardized test scores, etc.) to construct quantitatively objective profiles of the reality under study, focusing primarily on the context and impact of the TestEdge intervention.

By contrast, drawing on their rich pedagogical and therapeutic experience, the CGU team used a wide variety of formal and informal data collection and interpretative procedures (classroom observations, interviews, student drawings, etc.) anchored on the principles of authenticity, coherence, and trustworthiness as determinants of validity and reliability (Eisner, 1993; Denzin & Lincoln, 2005). The aim was to portray the actual social reality and experience of students and teachers in their classroom settings and schools. The expectation was that these representations would provide some insight and deeper understanding of what it means to be a teacher or a student under the mandatory order of standardized testing, and they would also provide a richer description of the meaning and impact of the intervention on both groups.

In short, the rich set of researcher perspectives and experiences and methodological mixes embedded in the quantitative and qualitative projects of TENDS were seen by the research team as an opportunity both to provide cross-data validity for evaluating the HeartMath tools as well as to directly address issues of implementation (Arguelles et al., 2006: 2).

[3] This section draws on the CGU Team's "Executive Summary Report: Inquiry Work Conducted by the CGU Team for the *TestEdge* National Demonstration Study" (Arguelles et al., 2006: 1-2).

Quantitative Instrumentation

Three self-report questionnaires were constructed to measure primary concepts: the Student Opinion Survey was completed by the students, while the Teacher Survey and the Classroom Relations Survey were both completed by the teachers (links to these instruments are provided for the electronic version of this report available at www.heartmath.org).

The Student Opinion Survey (SOS; Bradley & Atkinson, 2004b) contains 80 items and was constructed to measure student sociodemographic characteristics; a broad range of students' perceptions of their relationships and connections to teachers, peers, family, and school; positive and negative affect; emotional discord; ability to manage stress; and test anxiety. For the measurement of test anxiety, eight items from the Spielberger Test Anxiety Inventory (Spielberger, 1980), representing both the "Worry" and "Emotionality" constructs, were incorporated in the SOS.[4] The relationship between the constructs and items is given in Table V.3 (pages 70-72). In addition, the last page of the SOS contained an empty frame with the following prompt: "Please take a few minutes to draw or sketch a picture that shows how you feel when taking an important test (Use the space provided below)."

The SOS questionnaire was completed by students twice: first, early in January of 2005 to obtain a baseline measurement, and again two weeks after taking the final high-stakes exam (CST). In addition, the test anxiety section, which we titled the Test Opinion Survey (TOS) for the purposes of this study, was administered two additional times: one week before students took the CAHSEE in March and one week prior to the CST in April.

The Teacher Survey (TS; Bradley & Atkinson, 2004c) contains 68 items and was constructed to measure sociodemographic characteristics, teaching experience, workload, perceptions of school climate, opinions about standards-based testing, positive

[4] The Test Anxiety Inventory (TAI), developed by Charles Spielberger, is the most commonly used validated self-report instrument for measuring test anxiety and has been utilized in the majority of more recent studies of student test anxiety. The TAI provides a global measure of test anxiety as well as a separate measurement of two theoretically relevant components defined as "worry" and "emotionality." The "Worry" construct, which has been found to be most strongly correlated with depressed test performance in students with high test anxiety (Cizek & Burg, 2006, p. 17), is essentially a measurement of the psychological aspects of test anxiety (i.e., thought processes and emotions relating to the fear of testing and dread regarding the potential for negative evaluation or failure). The "Emotionality" construct provides a measure of the physical symptoms of test anxiety (e.g., nervousness, sweating, fidgeting, etc.).

and negative affect, and ability to manage stress. In order to compare students and teachers on the same constructs, most of the items in the TS were identical to those in the SOS. This questionnaire was completed by teachers twice: first, early in the school term (January 2005) to obtain a baseline measurement, and again in April 2005, two weeks after students took the final high-stakes exam.

The teachers also completed the Classroom Relations Survey (Bradley & Atkinson, 2004a) for each of their classes of students participating in the study. This questionnaire contains 28 items and was constructed to measure teachers' perception of the overall emotional tone and interactional dynamics that occur in a specific class, as well as the teaching strategies used in that class. These questionnaires were completed at the same times the Teacher Survey was completed.

In addition to the quantitative instruments described above, an 18-item multiple choice questionnaire—the TestEdge Assessment (Bradley, Goelitz, Atkinson, & Tomasino, 2004)—was developed to assess student comprehension of the lessons taught in the TestEdge program and to inquire into how students used the TestEdge tools in school, in their personal lives, and in test-taking situations. Students were asked to report their frequency practice and use of the tools in a variety of situations outside the classroom on a Likert scale.

The administration and psychometric integrity of the instrumentation was tested and verified in a pilot study conducted during the summer of 2004 at "Southern California High School" in Southern California, with 96 students in the 9th grade. Based on the results of a measurement validity and reliability analysis of the pilot data, some items were either deleted or modified to increase psychometric performance of the instrumentation. The classroom observational protocols were also tested at this site during the same time period. A detailed account of this pilot study is provided by Schroeder (2006).

To measure test performance, students' scores from the CAHSEE and the CST 2004 and the CST 2005 in English-Language Arts, Mathematics, and Science were obtained for both the intervention and control schools. CST 2004 was designated as the pre-intervention test score variable, and CST 2005 was designated as the post-intervention test score variable.

Physiological Measures

An electrophysiological sub-study was conducted in order to investigate the extent to which students had learned the techniques taught in the TestEdge program by measuring the degree to which they had developed the ability to shift into the coherence state prior to engaging in a stressful situations such as high-stakes testing. Utilizing objective physiological measures, this study collected pre- and post-intervention recordings of heart rate variability (HRV) from a random sample of students from both school sites, stratified by test anxiety level and gender. In this study, to simulate a stressful testing situation, students completed an experimental procedure that included a computerized version of the Stroop Test—a color-word conflict test (a standard protocol used to induce psychological stress)—while continuous HRV recordings were gathered through noninvasive measurement of the pulse. During the pre-intervention administration, students were asked to prepare themselves to take the test using whatever methods they typically used when preparing to perform a challenging test or activity. In the post-intervention session, students in the intervention group were instructed to use one of the positive emotion-focused coherence-building techniques they had learned in the TestEdge program to ready themselves for the test, while the control group students again used their own methods. From the inter-beat interval data, a number of indices of HRV and a measurement of heart rhythm coherence (the key marker of the psychophysiological coherence state) were derived.

Qualitative Instrumentation

To supplement the data gathered from the quantitative questionnaires, an Observational Protocol (Arguelles & Schroeder, 2004) was constructed to obtain measures of the socioaffective climate and interaction patterns in the classrooms and, more broadly, within the schools. These protocols included items from the Los Angeles County TPE Cooperating Teacher Assessment and from an instrument designed by Costa and Kallick (2000), both of which have proven effective in measuring teaching practices, student behaviors, and classroom climates in ethnically and socioeconomically diverse schools. The details of the observational protocols are provided by Schroeder (2006). School and classroom observations were conducted prior to the implementation of the TestEdge program, during the month of January 2005 (Time 1), again during the implementation of the program in March (Time 2), and finally in May (Time 3), after the TestEdge program was completed.

In addition, teacher and school administrator interviews were conducted at all sites to gather data on the teachers' and administrators' experience in implementing the TestEdge curriculum. The teacher and school administrator interview data were the primary data used to assess the experience of the implementation of the intervention, how the program was received by students, and any student benefits at the secondary sites.

Statistical Analysis

The statistical analysis was designed to verify the study's two primary hypotheses: 1) that the TestEdge intervention would reduce student test anxiety and increase test performance; and 2) that the intervention would also produce improvements in student attitudes, in student emotional stability and behaviors, and in the quality of classroom interactions.

Following data cleaning and verification, the questionnaire responses from students and teachers, observational data, and student test scores from the CAHSEE and CST standardized tests were entered into an SPSS database. The database was organized into a time series format involving four time points at which various measurements and observations were recorded during the study.

A univariate analysis using descriptive statistics (frequency distributions, means, standard deviations, standard errors, etc.) was conducted to check all variables for their technical integrity (case count, variability/distribution, outlying cases, etc.), and to build a descriptive profile of the data. Bivariate analysis, involving cross-tabulations and correlations, was conducted to verify respondent consistency on certain key variables and to identify data discrepancies and anomalies for investigation and resolution in order to construct a usable database. Bivariate procedures were also used to evaluate the achieved sample of students in relation to the sample selection criteria and to check sample representativeness and sample bias. An analysis of the validity and reliability of measurement was conducted on each of the multi-item scales included in the student survey and also on other multi-item constructs. Pearson correlations, Cronbach's alpha, and factor analysis were used to verify the measurement integrity of the scales and multi-item constructs that were employed in the main statistical analysis.

Analysis Strategy

Once we had completed an exploratory analysis of the data, statistical analysis was guided by the following strategy. We started with an aggregated analysis of the whole sample (i.e., students from both the experimental school and control school together). This provided a descriptive picture of the sample in terms of the primary constructs—test anxiety, the SOS scales, and test performance—which we contextualized with a breakdown by basic sociodemographic variables. We then investigated the nature and incidence of test anxiety within the 10^{th} grade student population and its relationship to test performance, using multivariate techniques such as analysis of covariance (ANCOVA), multiple regression, and discriminant function analysis.

In moving to the key question of intervention effects and the statistical comparison of the experimental school against the control school, we began with the question of the equivalence of the two groups. A key requirement of the quasi-experimental design is that the experimental and control group samples are equivalent, or closely matched, at the time of baseline measurement on all key variables. Any difference found on these factors raises the possibility that any changes observed in test anxiety or in test performance may be explained, at least in part, by this difference. We performed a single factor analysis of variance (ANOVA) to address this question, and then conducted a within-groups analysis, using factor analysis and ANCOVA, of each sample to understand the nature and magnitude of pre-post changes observed within each school system. This not only provides a clearer understanding of the distinctive endogenous conditions and forces at work in each school system, but the identification of any within-groups difference may also help explain the successes and failures of the TestEdge intervention in the experimental group.

We then turned to the important question of pre-post differences on test anxiety and test performance in the two samples, and the degree to which any observed differences were attributable to the expected effects of the TestEdge intervention. We used an ANCOVA, which adjusts for differences in means at baseline measurement, to compare the pre-post changes in test anxiety, test performance, and constructs of the SOS scales in the two groups. However, due to unanticipated sampling differences in socioeconomic level and ethnic composition between the two schools, only a small degree of between-groups change was observed in student test performance. When these differences are statistically controlled by using a matched-groups procedure, the expected effects of the intervention on the experimental group—a reduction in test anxiety in conjunction with an increase in test performance—were observed.

This entailed a change in the level of analysis from that of the total samples of the experimental and control groups, to one of an investigation of intervention effects by comparing various sub-samples within the two groups. We used two matched-group strategies. One involved the use of multivariate techniques to construct matched-group comparisons on various combinations of sociodemographic variables and the SOS scales. This enabled the employment of statistical controls in which a certain combination of variables is held constant in a given matched-group sub-sample pair while the intervention effects on test anxiety and test performance are determined by comparing the experimental group sub-sample with the corresponding control group sub-sample. This approach revealed a number of findings showing evidence of a successful intervention effect in a number of sub-populations within the experimental group. For these sub-samples, the final step was to investigate the expected relationship between various aspects of emotional disposition (as measured by the SOS scales) and the changes in test anxiety and test performance.

For the second approach, matched groups of classes were constructed by matching classes by mean class test performance at baseline measurement, on the 9th grade CST. This procedure resulted in the construction of two different sets of matched-class groupings: one set of three groupings matched on the 9th grade mean class English-Language Arts test score, and a second set of two groupings matched on the 9th grade mean class Mathematics test score. We then conducted a series of ANCOVAs to investigate the relationship between various aspects of emotional disposition (as measured by the SOS scales) and the changes in test anxiety and test performance in each sub-sample experimental group–control group pairing.

Before turning our attention to the physiological study, we grouped the data by the students' homeroom class to investigate the question of teacher effects. We also examined the degree to which there was evidence of a relationship between changes in test anxiety, emotional disposition, and test performance of individual students. This included the contextual effects of the social group—the class—to which students are exposed through relations with their teachers and classmates. As a part of this analysis, we also investigated the relationship between student perceptions and teacher perceptions in an analysis of the common items on the SOS and the Teacher Survey.

Statistical analysis of the physiological study was conducted in three phases. The first entailed an analysis of the students in the physiology study sample to evaluate the degree to which the two groups were equivalent in terms of sociodemographic characteristics and baseline measures of test anxiety, test performance, emotional disposition (as indicated by the SOS scales), and electrophysiological measures of HRV. The second phase involved analyzing pre-post changes in the electrophysiological data during the resting period and in the stress preparation period of the Stroop Test experiment. An ANCOVA was used to compare the pre-post changes in HRV measures in the two groups of students between the resting period and the stress preparation period, to test the hypothesis that students in the experimental group, following the intervention, would move to a state of increased heart rhythm coherence during the stress preparation period. The third phase of the analysis involved integrating the physiological data into the main database and conducting a series of ANCOVAs to investigate the degree to which the differences between the experimental group and the control group on the pre-post changes in test anxiety, test performance, and HRV measures were consistent with the expected effects of the TestEdge intervention. Matched-group comparisons of sub-samples within the physiological study population were also analyzed. This was done for two reasons: first to identify the distinguishing sociodemographic characteristics of groups of students for whom the expected effects of the intervention were evident; and, second, to build a deeper understanding of the more complex relationships among test anxiety, psychophysiological coherence, and test performance, as revealed by the data from the physiological study. Finally, we used discriminant function analysis in an effort to determine the degree to which the experimental and control groups were differentiated from each other by a common statistical function comprised of post-intervention changes in test anxiety, physiological coherence, and test performance.

Statistical Inference

Studies which collect large volumes of quantitative data, like this one, and run hundreds of statistical analyses to examine relationships among many hundreds of variables, face a number of basic issues regarding statistical inference. Among the most important questions of statistical inference is that involving the commission of Type 1 and Type 2 errors of inference: accepting statistical evidence as proof of the veracity of a hypothesis when, in fact, the hypothesis is invalid; or rejecting the validity of a hypothesis on the basis of the statistical evidence at hand when, in fact, the hypothesis is true. While it is virtually impossible to ensure that not a single Type 1 or Type 2 error has been committed in interpreting the enormous volume of statistical results generated in the course of the statistical analysis conducted in this research, we were guided by the following considerations in our effort to keep the commission of these errors of inference to a minimum.

A major area of inferential vulnerability lies with results for which the confidence interval falls within the reach of chance: that is, results which are just under or just over the threshold of chance, say within a few hundredths of $p = 0.05$. We believe inferential vulnerability is greatest when interpreting what appears to be a single result that does not appear to be part of an existent pattern of evidence. We used three principles to guide our interpretive inference in deciding whether to accept or reject a given statistical result as valid. A result of marginal statistical significance must meet either of the first two principles; for a significant result involving small case counts, the third principle advises interpreting it with caution:

- The first was that the result must be consistent with prior research and the study's major theoretical expectations. This principle would rule out a marginal result that did not make sense within the study's theoretical interpretive framework.
- The second principle was that the result must be consistent with an existent pattern of findings or must be corroborated by data from another source or method. In essence, this is the requirement of internal consistency, which holds that there must be some pattern of evidentiary support for the result, beyond that of the result itself.
- The third principle applied to a result from analysis involving small case counts, and required that even if the confidence interval for the result was well beyond the 0.05 level, the result must be treated with caution. A primary source of inferential vulnerability is due to the sampling error involved in the multivariate selection procedures used to construct a subgroup with relatively few cases.

A second major issue concerns the question of the generalizability of the study's primary findings. This can be addressed in two ways. The first approach is an evaluation of the degree to which the study populations—the samples of students from the two high schools—are representative of other high school populations in California and also across the nation. And a related question for evaluation concerns the degree to which the conditions encountered and procedures used to implement the TestEdge program in this study are the same as those typically experienced when TestEdge is used in American schools. Both questions are addressed in Chapter XV, "Discussion of Major Findings."

Causal Inference

Evaluating the evidence collected in this study to determine whether or not the TestEdge intervention was successful involves causal inference: assessing the degree to which, all other things being equal, the TestEdge intervention actually caused changes in the internal processing and behavior of students in the experimental school that resulted in the observed outcomes. To address this question of causality, the study utilized a quasi-experimental design in which one school was treated as the treatment group—into which the TestEdge program was introduced—and the second school was treated as a control.

Operationally, the experimental design required pre and post measurements on each side of the TestEdge intervention: one measurement at baseline before the TestEdge intervention was implemented, and a second measurement at the end of the treatment period, just before the students took the CST. In essence, this means that the study utilized a two-panel administration of the instrumentation—repeated administration of the same discrete measurements at two different time points.[5] While this research design meets Campbell and Stanley's (1963) criteria for causal inference in a quasi-experimental design, there are some additional considerations that require the exercise of caution when weighing evidence of causality in studies, like this one, that utilize a two-panel repeated measures design. These considerations and caveats are examined in Chapter XV where we weigh the significance of the major findings in relation to the study's limitations.

[5] It should be noted that the test anxiety instrument was administered in a four panel format. In addition to pre and post measurement, the test anxiety instrument was administered two additional times: one week before students took the CAHSEE in March and one week prior to the CST in April.

Section Two

Results of the Primary Study: Assessing the Effectiveness of the TestEdge Intervention

Chapter V
Implementation of the Primary Study

The primary study utilized a multi-methods, randomized approach with experimental and control groups. Employing both quantitative and qualitative measures, this study was designed to rigorously investigate the relationship between the intervention and resultant changes in student psychophysiological states, attitudes, aspirations, social and classroom behaviors, test anxiety, and test performance.

Sample Selection

For the primary study, high schools in eight states were considered for inclusion in the study's sample. After a careful review of the available data in relation to the study design requirements, prospective schools for inclusion in the study were narrowed to high schools in three states—Ohio, Florida, and California. Efforts to negotiate timely participation in the study were unsuccessful in the Florida schools. For similar reasons, it was also difficult to find suitable schools in Ohio. Looking closer to home, we identified two schools in close proximity to the HeartMath Research Center. The schools identified for the primary study were "Valley High School," located in the North Central Coast, and "Bay High School," located in the San Francisco Bay Area.

For the two California high schools, final determination of study status (experimental versus control group) was made by using a computerized random selection program. The result was selection of "Valley High School" as the experimental group, which meant that "Bay High School" was assigned to be the (wait-list) control group. Each school was then contacted and informed of the outcome of the random selection process, and a meeting to familiarize the principals with the operational implications of participation in the study was arranged and attended by the Institute of HeartMath's senior research staff.

Despite our efforts to locate and solicit the participation of two schools that were comparable on sociodemographic characteristics and academic level, a number of important differences were apparent both from observations of community context and also from survey data collected from the student respondents in the two schools.

School Context[6]

Observations of the context for each school site were guided by a community profile protocol (Wolk & Behr, 2004) that was developed utilizing publicly available data accessed electronically. The intervention school is located on an older campus (built in 1875, remodeled in 1980) surrounded by a predominantly Hispanic low-to-middle income community. It is a large school of approximately 3,000 students (2,915 enrolled in the 2004–2005 academic year, of which 720 were enrolled in the 10th grade). Most of the students are either Hispanic (53.4%) or White (42.4%). In 2004–2005, students whose primary language was other than English made up 30.7% of the school's population, and 27.6% qualified for free or reduced price meals. The school's API fell from 672 in 2003 to 666 in 2004, thus failing to meet the API growth target for the academic year 2003–2004.

By comparison, the control site is located in a newer school (built in 1952, remodeled in 1998) in a middle and upper-middle income community. The school has a smaller student population of approximately 2,000 students (2,041 were enrolled in the 2004–2005 academic year) of which 490 were in the 10th grade. Slightly more than half are White (54.2%), with significant numbers of Hispanic (21.3%) and Asian (11.3%) students. Students whose primary language was other then English comprised 19.9% of the student population, and 13.4% qualified for free or reduced-price meals. In contrast to the intervention school, the control school's API rose by 31 points, from 709 to 740, and was successful in making the 5% improvement toward the interim API goal of 800 in the 2003–2004 academic year. While the two schools were closer

[6] This section is drawn in part from Arguelles et al. (2006: 3-4).

in API ranking at the time of sample selection, with a difference of 37 points according to the 2003 data (intervention school = 672 and control school = 709), the substantial improvement in API in the control school from 2003 to 2004, just noted, has doubled the difference between the schools to 74 points.

Teacher workloads were higher and per pupil expenditures were lower in the intervention site compared to the control site. The intervention and control schools also differed in terms of stated goals. The intervention site school's goals were mostly anchored on a philosophy of discipline that would lead to the creation of a safe school and warm and friendly classroom environments. The control site school emphasized goals that were mostly conducive to students fulfilling requirements necessary for admission to colleges and universities, and therefore stressed the importance of high stakes testing and the like.

There are also differences in community context, as documented by Hollingsworth (2007, see Table 4). When compared to the control school, the intervention school is located in a community with a larger population (344,413 versus 25,123) that had a notably younger age structure (median age: 29 versus 39 years), less formal education (Bachelor Degree: 10.8% versus 30.5%), higher unemployment (5.1 versus 2.0), and less affluent households (gross value: \$266,300 versus \$593,200).

Overall, it is clear that, despite the effort to select comparable schools, there are notable differences between the two schools in terms of student and community ethnic composition and socioeconomic status and built environment, size, goals, teacher workload, per pupil expenditures, and API ranking. Moreover, the unanticipated large improvement in the control school's API ranking from 2003 to 2004 further exacerbated the lack of match between the two schools on the key dependent variable of academic performance. If anything, these differences and this change in academic ranking work against successful demonstration of the TestEdge program in the intervention school, as will be discussed momentarily.

Sampling Results

Table V.1 presents the results of implementing the sampling plan to 10th graders in the two schools. The total sample of 10th graders reported as enrolled in the two schools was 1,210, of which 720 (59.5%) were in the intervention school (experimental group) and 490 (40.5%) were in the control school. Of this total, 230 students (19.0%) were ineligible for the study, either because parental permission for participation was denied (37 students; 3.1%), because the parental permission forms were not returned (53 students; 4.4%), or because students were unavailable for participation in the study due to conflicts with their class schedule (140 students; 11.6%). This reduced the total number of students eligible for participation in the study to 980, 636 (64.9%) of which were in the experimental group and 344 (35.1%) in the control group. The greater representation of the experimental group over the control group, when the total sample proportions are compared with the breakdown of the study population of eligible students, is explained by the greater proportion of students in the control group with class schedule conflicts (42 versus 98 students, respectively).

Table V.1 Sample Characteristics

	Experimental Group			Control Group		
	Students (N)	**Enrolled (%)**	**Eligible (%)**	**Students (N)**	**Enrolled (%)**	**Eligible (%)**
Reported Enrollment	720	100%		490	100%	
Parental permission denied	-15	-2%		-22	-4%	
Permission forms not returned	-27	-4%		-26	-5%	
Class schedule conflicts	-42	-6%		-98	-20%	
Pre-Study Population (Eligible Students)	636	88%	100%	344	70%	100%
Absent or no longer in class	-28	-4%	-4%	-6	-1%	-2%
Students providing unusable data	-6	-1%	-1%	-6	-1%	-2%
Usable Pre-Study Surveys	602	84%	95%	332	68%	97%
Post-Study Population (Eligible Students)	602	84%	95%	332	68%	97%
Absent or no longer in class	-64	-9%	-10%	-27	-6%	-8%
Missing/lost forms	0	0%	0%	-12	-2%	-3%
Students providing unusable data	-18	-3%	-3%	-18	-4%	-5%
Surveys from resource level classes	-32	-4%	-5%	-14	-3%	-4%
Usable Sample	488	68%	77%	261	53%	76%

During administration of the study's instruments, additional losses of student participants occurred in both the pre- and post-study populations of eligible students. For the pre-study population as a whole, 46 students (4.6%) were either absent or no longer in class, or provided unusable data during baseline measurement. This resulted in a usable sample of 934 cases, which is 77.2% of the total sample or 95.3% of the pre-study population of eligible students. When divided by intervention status, these losses reduced the usable samples to 602 cases in the experimental group and 332 cases in the control group. These usable samples represent 83.6% and 67.8%, respectively, of the experimental and control groups when computed as a proportion of the total sample, or 94.7% and 96.5%, respectively, when computed as a proportion of the pre-study population of eligible students.

By the end of the school quarter, further losses of student participants were incurred, which reduced the post-study population of eligible students from 934 to 749 cases. For the study population as a whole, 185 (19.8%) students were either absent or no longer in class, had missing or lost questionnaires, provided unusable data, or returned surveys from resource level classes. The pre-study reported enrollment of 1,210 represents 61.9% with usable data in the post-study sample; the 934 pre-study population of eligible students represents 80.0% with usable data. Broken out by intervention status, the losses by the post-study period reduced the usable sample to 488 cases in the experimental group and 261 cases in the control group. These numbers represent 67.8% and 53.1%, respectively, of the original total enrolled populations in the experimental and control groups, and 81.1% and 78.6%, respectively, of the post-study population of eligible students.

Sample Characteristics and Representativeness

Table V.2 presents data on selected variables to assess the degree to which the usable sample is representative of the total reported enrolled 10th grade populations in each of the two schools. These data also give some indication of the comparability of the two schools.

The data in Table V.2 are for usable cases in the post-study samples. They show that these samples are closely matched with the 10th grade total enrolled population in relation to age, gender, ethnicity, and family structure in both the experimental and control group schools.

Table V.2 Sample Representativeness by Selected Characteristics

	Entire sample N = 749	Experimental N = 488	Control N = 261
Age, years (mean, sd)	15.2, 0.43	15.3, 0.44	15.2, 0.42
Gender			
Male	47%	52%	40%
Female	53%	48%	60%
Ethnicity			
Caucasian	43%	37%	54%
Hispanic	36%	50%	12%
Asian	9%	3%	20%
Other	8%	7%	10%
African American	1%	1%	2%
Pacific Islander	1%	1%	1%
American Indian	1%	1%	1%
Parents			
Both biological parents	64%	63%	66%
Single bio parent	13%	13%	12%
Mixed family, one bio parent	10%	12%	7%
Dual custody	8%	8%	10%
Other	4%	3%	4%
Relatives	2%	1%	2%

A comparison of the samples from the two schools shows that while there is broad similarity on most of the sociodemographic characteristics, a number of differences are evident in the usable post-study samples. The biggest difference involves ethnicity—the strikingly greater proportion of Hispanic students in the experimental group (50% versus 12%) and the correspondingly higher proportion of Caucasian students (54% versus 37%) and Asian students (20% versus 3%) in the control group. A somewhat smaller difference is also apparent for gender—almost evenly split between males and females in the experimental group (52% and 48%, respectively), but favoring females by a twenty-percent margin in the control group (60% versus 40%, respectively). Further data on the comparability of the two schools are presented below.

Sample Bias and Potential Impact on Findings

Despite the efforts to select comparable schools, it is clear from the community profile data and the data on the student samples that the two schools were not well matched on ethnic composition or socioeconomic status, and given the control school's improved API standing from the 2003-04 to 2004-05 academic year, were also not closely matched academically. While this lack of equivalence can be overcome, to some degree, by multivariate analysis techniques, it nonetheless constitutes a significant contextual difference between the two schools. Moreover, it is unlikely that the use of multivariate procedures would successfully control for the *qualitative* differences between the schools as whole systems, manifest in such factors as teacher workload, school resources, and educational goals and philosophy.

Lower teacher workloads, greater resources, more academically focused educational goals and philosophy, and more affluent and educated community contexts are all factors likely to promote the creation of a more effective environment for academic learning and performance in schools. These factors all favor the control school over the intervention school and place the latter at a disadvantage as a less effective environment for academic learning and performance. Under these conditions, therefore, in order to demonstrate a TestEdge effect, the intervention school must first overcome the effects of these relative disadvantages on socioemotional measures, test anxiety, and test performance, and additionally show the intervention effect as an otherwise unexplained residual difference on these measures. In short, these differences between the two schools likely made successful demonstration of the efficacy of the TestEdge intervention significantly more difficult. More will be said about this later, when presenting the results and discussing the main findings.

Measurement and Operationalization

The quantitative analysis that follows for the Primary Study is based on data derived from four sources (described in greater detail in the instrumentation section, above). The first was the Student Opinion Survey questionnaire,[7] the second was the Teacher Survey questionnaire,[8] the third was the classroom interaction observation protocol,[9] and the fourth was student scores on the CAHSEE and the CST. While certain questions and data items were operationalized in a single-variable measurement format, most questions and data items were constructed into multivariate scales for

[7] A copy of the Student Opinion Survey questionnaire is provided in Appendix 3.
[8] A copy of the Teacher Survey questionnaire can be obtained by contacting HeartMath.
[9] A copy of the classroom interaction observation protocol can be obtained by contacting HeartMath.

more robust measurement. For clarification, Table V.3 presents the relationship between each concept and measurement construct for all variables and scales used in this analysis.

Table V.3 Operationalization of Constructs from Items in the Student Opinion Survey

Feelings about School

__**Answer the following questions about your school**__

6. I feel close to people at this school
7. I am happy to be at this school
8. I feel safe in this school

Teacher Support

__**At my school, there is a teacher or some other adult…**__

9. who really cares about me
10. who listens to me when I have something to say
11. who believes that I will be a success

Educational Plans

__**How true do you feel these statements are about you personally?**__

12. I plan to graduate from high school
13. I plan to go to college or some other school after high school
14. There is a purpose to my life

Life Preparedness

__**I feel what I have learned at this school**__

15. will help me be successful as an adult
16. has prepared me to do well at college or get a good job
17. has inspired me to want more for myself out of life

Parental Support

__**In my home, there is a parent or some other adult …**__

18. who is interested in my schoolwork
19. who believes that I will be a success
20. who talks with me about my problems
21. who loves me no matter what

Positive Class Experience

__**Read the statements below and indicate how you feel while you are in this class**__

22. I enjoy this class and find it fun
23. I am pleased with how much I am learning
24. I feel that there are mostly good feelings among all of us in this class
25. I feel the teacher cares about me and my classmates as individuals

Extent of Friendship

__**For those classmates or friends you feel closest to, do you:**__

29. Do homework together
30. Feel very comfortable with each other
31. Talk on the phone
32. Confide in each other about personal situations or problems
33. Spend time alone together as friends
34. Have strong feelings about each other
35. Care a lot about each other

Table V.3 continued…

Positive Affect

In thinking about my life over the past few weeks, I generally find that:

36. I feel happy
39. I feel calm
41. I feel appreciative
42. I feel cheerful
45. I feel enthusiastic
46. I feel hopeful
48. I feel loved

Negative Affect

In thinking about my life over the past few weeks, I generally find that:

37. I feel stressed
38. I feel lonely
43. I feel sad
44. I feel angry
47. I feel depressed
49. I feel disappointed

Emotional Discord

In thinking about my feelings over the past few week, I generally find that

50. I feel powerless over what I am feeling
51. I don't always know clearly how I feel
52. I have opposite feelings from one moment to the next
53. I keep negative feelings bottled up inside
54. I feel overwhelmed by my feelings

Interactional Difficulty

In my interactions with others, there are situations in which:

55. I find it difficult to know what others are feeling
56. I have difficulty sharing my feelings with others
57. I don't feel that I am being heard and understood
58. I feel that I don't matter
59. I get into arguments or fights

Stress Management

When I am stressed out, I:

60. Try to manage my stress in the moment it happens
63. Try to breathe deeply or relax to calm myself down
64. Try to activate a positive emotion
66. Find it helpful to do physical exercise or some other activity

To handle my reactions to stress:

67. I try to change to positive feelings when I've had negative feelings for a while
68. I am able to calm myself down when feeling distressed
69. I am able to control my feelings effectively when feeling overwhelmed
70. I try to be less judgmental when feeling angry at myself
71. I try to change to hopeful feelings when feeling depressed
72. I try to become more peaceful when I'm feeling anxious

Table V.3 continued…

Test Anxiety Inventory
Following are a number of statements that people have used to describe their feelings about tests.
<u>Indicate how you generally feel:</u>

Test Anxiety-Global

73. I freeze up on important exams
74. The harder I work at taking a test, the more confused I get
75. During tests I find myself thinking about consequences of failing
76. During examinations I get so nervous that I forget facts I really know
77. I feel very jittery when taking an important test
78. During tests I feel very tense
79. I feel very panicky when I take an important test
80. I worry a great deal before taking an important examination

Test Anxiety Inventory - Worry sub-scale

73. I freeze up on important exams
74. The harder I work at taking a test, the more confused I get
75. During tests I find myself thinking about consequences of failing
76. During examinations I get so nervous that I forget facts I really know

Test Anxiety Inventory - Emotional sub-scale

77. I feel very jittery when taking an important test
78. During tests I feel very tense
79. I feel very panicky when I take an important test
80. I worry a great deal before taking an important examination

In terms of the fifteen scales developed from the SOS, Table V.3 shows that seven scales were constructs that measured various aspects of the student's school experience, life hopes, and home experience:

- Feelings about School
- Teacher Support
- Educational Plans
- Life Preparedness
- Parental Support
- Positive Class Experience
- Extent of Friendship

The remaining eight scales were constructs that measured various aspects of the students' feelings and emotional life, social relations, how they handle stress, and test anxiety:

- Positive Affect
- Negative Affect
- Emotional Discord
- Interactional Difficulty
- Stress Management
- Test Anxiety-Global
- Test Anxiety-Worry subscale
- Test Anxiety-Emotionality subscale

Scale Construction

A primary focus of the quantitative analysis was on the data from SOS. Table V.5 shows that with the exception of the five questions gathering background information in Section 1, questions 26 through 28 on the number of close friends, and the sketch at the end of the questionnaire, the remaining 76 items were scaled into fifteen multivariate constructs using the procedure described below. It will be recalled that the SOS scales were constructed from individual items with a Likert-scale response format. With one exception, all of the SOS scales have a 4-point ordinal scale metric; the exception is Feelings about School, which used a 5-point ordinal scale metric.

Starting with the grouping of questions into distinct sets of items in the questionnaire (e.g., "… questions about your school:" items 6–8; "At my school, there is a teacher or some other adult …:" items 9–11; etc.), an analysis of measurement reliability and validity was conducted using Cronbach's alpha and factor analysis to evaluate the degree to which the items in each set were internally consistent and measured the same underlying construct.

As the alpha coefficients show in Table V.4, twelve of the fifteen scale constructs achieved or exceeded the technical criterion for measurement adequacy—alpha coefficient ≥ 0.80; for these constructs, the alpha coefficient ranged from a low alpha value of 0.80 for Life Preparedness, Positive Class Experience, and Emotional Discord, to a high alpha value of 0.92 for the Test Anxiety-Global scale. For the other three constructs, the alpha coefficient was 0.72 for Interactional Difficulty, 0.62 for Feelings About School, and 0.47 for Educational Plans. While the construct for Educational Plans is included in the analyses that follow, it will be removed from all future publications since the alpha coefficient is well beneath the threshold of technical adequacy for internal consistency.

Table V.4 Cronbach Alpha Results

Student Opinion Survey Scale Reliability	
Feelings about School	0.62
Teacher Support	0.84
Educational Plans	0.47
Life Preparedness	0.80
Parental Support	0.81
Positive Class Experience	0.80
Extent of Friendship	0.86
Positive Affect	0.82
Negative Affect	0.86
Emotional Discord	0.80
Interactional Difficulty	0.72
Stress Management	0.85
Test Anxiety-Global	0.92
Test Anxiety-Worry	0.87
Test Anxiety-Emotional	0.90

To statistically confirm construct validity—that the assignment of items to each construct was statistically optimal—a factor analysis with varimax rotation was performed into which all 76 items were entered. The results, reported in Table V.5 (next page), show that, with only a few exceptions, the original classification of items into their assigned constructs was optimal. These exceptions were carefully examined in terms of both item content and the factor loading correlation coefficients. Because the factor loadings were close in value to those items in the original classification, and because the item content better fit the content theme of those in the original classification, we made the decision to keep all items in their original classification, as grouped in the SOS (see Table V.3 on previous pages).

Table V.5 Factor Analysis Results for the SOS Scales

Q#	Final Scale	Item	
47	Negative Affect	I feel depressed	**0.76**
43	Negative Affect	I feel sad	**0.75**
49	Negative Affect	I feel disappointed	**0.72**
37	Negative Affect	I feel stressed	**0.68**
44	Negative Affect	I feel angry	**0.66**
38	Negative Affect	I feel lonely	**0.58**
36	Positive Affect	I feel happy	**-0.55**
58	Interactional Difficulty	I feel that I don't matter	**0.54**
59	Interactional Difficulty	I get into arguments or fights	**0.52**
54	Emotional Discord	I feel overwhelmed by my feelings	**0.51**
39	Positive Affect	I feel calm	**-0.44**
53	Emotional Discord	I keep negative feelings bottled up inside	**0.40**
79	Test Anxiety-Global	I feel very panicky when I take an important test	**0.84**
78	Test Anxiety-Global	During tests I feel very tense	**0.83**
73	Test Anxiety-Global	I freeze up on important exams	**0.82**
76	Test Anxiety-Global	During examinations I get so nervous that I forget facts I really know	**0.82**
77	Test Anxiety-Global	I feel very jittery when taking an important test	**0.79**
75	Test Anxiety-Global	During tests I find myself thinking about consequences of failing	**0.77**
80	Test Anxiety-Global	I worry a great deal before taking an important examination	**0.75**
74	Test Anxiety-Global	The harder I work at taking a test, the more confused I get	**0.70**
71	Stress Management	I try to change to hopeful feelings when feeling depressed	**0.74**
72	Stress Management	I try to become more peaceful when I'm feeling anxious	**0.73**
68	Stress Management	I am able to calm myself down when feeling distressed	**0.70**
70	Stress Management	I try to be less judgmental when feeling angry at myself	**0.68**
67	Stress Management	I try to change to positive feelings when I've had negative feelings for a while	**0.65**
64	Stress Management	Try to activate a positive emotion	**0.62**
69	Stress Management	I am able to control my feelings effectively when feeling overwhelmed	**0.59**
60	Stress Management	Try to manage my stress in the moment it happens	**0.52**
63	Stress Management	Try to breathe deeply or relax to calm myself down	**0.44**
33	Extent of Friendship	Spend time alone together as friends	**0.81**
34	Extent of Friendship	Have strong feelings about each other	**0.80**
35	Extent of Friendship	Care a lot about each other	**0.80**
32	Extent of Friendship	Confide in each other about personal situations or problems	**0.79**
31	Extent of Friendship	Talk on the phone	**0.70**
30	Extent of Friendship	Feel very comfortable with each other	**0.67**
29	Extent of Friendship	Do homework together	**0.47**

Continued
Table V.5 Factor Analysis Results for the SOS Scales

18	Parental Support	who is interested in my schoolwork	**0.79**
19	Parental Support	who believes that I will be a success	**0.79**
21	Parental Support	who loves me no matter what	**0.77**
20	Parental Support	who talks with me about my problems	**0.74**
22	Positive Class Experience	I enjoy this class and find it fun	**0.79**
23	Positive Class Experience	I am pleased with how much I am learning	**0.75**
25	Positive Class Experience	I feel the teacher cares about me and my classmates as individuals	**0.70**
24	Positive Class Experience	I feel that there are mostly good feelings among all of us in this class	**0.69**
15	Life Preparedness	will help me be successful as an adult	**0.75**
16	Life Preparedness	has prepared me to do well at college or get a good job	**0.73**
17	Life Preparedness	has inspired me to want more for myself out of life	**0.72**
45	Positive Affect	I feel enthusiastic	**0.64**
42	Positive Affect	I feel cheerful	**0.56**
48	Positive Affect	I feel loved	**0.53**
46	Positive Affect	I feel hopeful	**0.47**
41	Positive Affect	I feel appreciative	**0.46**
40	#N/A	I feel anxious	**0.41**
10	Teacher Support	who listens to me when I have something to say	**0.80**
9	Teacher Support	who really cares about me	**0.79**
11	Teacher Support	who believes that I will be a success	**0.78**
27	#N/A	# of close friends in this school	**0.87**
28	#N/A	# of close friends in the community	**0.80**
26	#N/A	# of close friends in this class	**0.75**
51	Emotional Discord	I don't always know clearly how I feel	**0.67**
52	Emotional Discord	I have opposite feelings from one moment to the next	**0.66**
50	Emotional Discord	I feel powerless over what I am feeling	**0.52**
56	Interactional Difficulty	I have difficulty sharing my feelings with others	**0.74**
57	Interactional Difficulty	I don't feel that I am being heard and understood	**0.52**
55	Interactional Difficulty	I find it difficult to know what others are feeling	**0.50**
61	#N/A	Talk to a friend, family member, or teacher	**-0.31**
12	Educational Plans	I plan to graduate from high school	**0.74**
13	Educational Plans	I plan to go to college or some other school after high school	**0.65**
62	#N/A	Watch TV or get on the Internet to distract myself from my feelings	**0.67**
65	#N/A	Find I want to sleep more than usual	**0.51**
66	Stress Management	Find it helpful to do physical exercise or some other activity	**0.46**
8	Feelings about School	I feel safe in this school	**0.59**
7	Feelings about School	I am happy to be at this school	**0.54**
6	Feelings about School	I feel close to people at this school	**0.35**
14	Educational Plans	There is a purpose to my life	**-0.44**

Extraction Method: Principal Component Analysis. | Rotation Method: Varimax with Kaiser Normalization.

Chapter VI
Analysis by Whole Sample

Before presenting the results of the HeartMath intervention, we begin with an analysis of the whole sample. This provides a descriptive picture of the sample in terms of the primary constructs—test anxiety and test performance—and a view of the sample in relation to the SOS scales. Following a breakdown of these factors by the basic sociodemographic variables, we present the results of an analysis of the relationship between test anxiety and test performance. This will set the stage for a series of analyses aimed to determine and evaluate the intervention effects which occupy the remainder of this report.

Baseline Descriptive Analysis

We begin this section with a pre-study baseline description of the whole sample in terms of student response on the primary thirteen SOS scales and their CST 9th grade test performance using univariate statistics (see Tables VI.1 and VI.2).

Beginning with the 9th grade California Standards Test scores, which were used as the baseline measures of test performance, the mean score overall was notably higher for English-Language Arts (360 points) than it was for Science (338 points) and Mathematics (333 points) (see Table VI.1). In reviewing the remaining data in Table VI.1, the reader is reminded that with the exception of the 5-point ordinal metric for the Feelings about School scale, the SOS scales have a 4-point ordinal scale metric. Bearing this in mind, the mean level of global test anxiety was 2.39, a little less than midway between the "Sometimes" and "Often" categories. The same level of mean test anxiety was also observed on each of the subscales—2.39 for the Worry subscale and 2.40 for the Emotionality subscale.

The data in Table VI.1 show that, on average, students were most positive about their Educational Plans (Mean = 3.77), Feelings about School (3.63), and Parental Support (3.51). They also felt fairly positive about Life Preparedness (3.06), Extent of Friendship (2.98), Teacher Support (2.93), Positive Class Experience (2.92), and Positive Affect (2.78). Their mean response for Stress Management (2.37) was somewhat lower, and lower still were their reports of experiencing Emotional Discord (2.17), feeling Negative Affect (2.13), and having Interactional Difficulty (1.93).

Table VI.1 Student Opinion Survey and CST Baseline Means

	Entire sample (N=749)					
	Mean	SEM	SD	Min	Max	Range
CST English - Language Arts 9	360	1.97	52.39	234	534	300
CST Math 9	333	2.03	53	218	544	326
CST Science 9	338	1.66	42	225	494	269
Test Anxiety-Global	2.39	0.03	0.89	1	4	3
Test Anxiety-Worry	2.39	0.03	0.92	1	4	3
Test Anxiety-Emotional	2.40	0.04	0.96	1	4	3
Feelings about School	3.63	0.03	0.71	1	5	4
Teacher Support	2.93	0.03	0.79	1	4	3
Educational Plans	3.77	0.01	0.39	2	4	2
Life Preparedness	3.06	0.03	0.70	1	4	3
Parental Support	3.51	0.02	0.65	1	4	3
Positive Class Experience	2.92	0.03	0.68	1	4	3
Extent of Friendship	2.98	0.03	0.69	1	4	3
Positive Affect	2.78	0.02	0.62	1	4	3
Negative Affect	2.13	0.03	0.70	1	4	3
Emotional Discord	2.17	0.03	0.73	1	4	3
Interactional Difficulty	1.93	0.02	0.61	1	4	3
Stress Management	2.37	0.02	0.62	1	4	3

A finer-grained picture of the pattern of student response on the SOS scales is given in Table VI.2, which provides the breakdown of quartile means for each scale. Students in all quartiles are highly optimistic about their future educational plans and life; in that quartile means range from 3.22 (1st quartile) to 4.00 (both 3rd and 4th quartiles. They are almost as positive about Parental Support, for which the 2nd, 3rd, and 4th quartile means are 3.58, 3.96, and 4.00, respectively. Based on a 5-point scale, their Feelings About School are also quite positive, with 2nd, 3rd, and 4th quartile means of 3.53, 3.89, and 4.41, respectively. On Teacher Support, Life Preparedness, Extent of

Friendship, and Positive Affect, the mean response (based on a 4-point scale) over the top three quartiles is still quite high (2.73–3.88, 2.89–3.89, 2.76–3.80, 2.82–3.74, and 2.57–3.55 from the 2nd to 4th quartiles, respectively).

However, there is also evidence of frequent experience of test anxiety (Global) for a significant portion of the sample in that the means for the 3rd and 4th quartiles were 2.71 and 3.60, respectively. A similar picture is revealed for the experience of Negative Affect and Emotional Discord: the means for the 3rd and 4th quartiles were 2.26 and 3.12, and 2.35 and 3.18, respectively. On the other hand, the frequency of Interactional Difficulty is noticeably lower by quartile, ranging from a mean of 1.26 to 2.77 from the 1st to the 4th quartiles. Finally, while the 1st and 2nd quartiles appear less effective in Stress Management (1.60 and 2.15, respectively), the 3rd and especially the 4th quartiles appear noticeably better (2.56 and 3.19).

Table VI.2 Baseline SOS Quartile Means

	1st	2nd	3rd	4th
Test Anxiety-Global	1.29	2.00	2.71	3.60
Test Anxiety-Worry	1.24	1.97	2.71	3.66
Test Anxiety-Emotional	1.22	1.94	2.75	3.72
Feelings about School	2.68	3.53	3.89	4.41
Teacher Support	1.86	2.73	3.24	3.88
Educational Plans	3.22	3.88	4.00	4.00
Life Preparedness	2.11	2.89	3.35	3.89
Parental Support	2.59	3.58	3.96	4.00
Positive Class Experience	2.03	2.76	3.20	3.80
Extent of Friendship	2.02	2.82	3.35	3.74
Positive Affect	1.96	2.57	3.01	3.55
Negative Affect	1.34	1.81	2.26	3.12
Emotional Discord	1.30	1.85	2.35	3.18
Interactional Difficulty	1.26	1.65	2.02	2.77
Stress Management	1.60	2.15	2.56	3.19

Table VI.3 presents the baseline test anxiety for the whole sample and its breakdown by gender, ethnicity, family composition, and class academic level. Taking student responses to the Test Anxiety-Global scale and dividing them into Low test anxiety (TAI-Global scale score of <2.0, less than "Sometimes"), Medium test anxiety ($\geq$2.1 - $\leq$3.0, more than "Sometimes" but less than "Often"), and High test anxiety (>3.0, "Often" or "Almost Always"), the table shows that while almost forty-percent (39.3%) had a low level of test anxiety, more than sixty-percent (60.7%) had medium (34.5%) or high (26.2%) levels of test anxiety. Moreover, although there is little evidence of a relationship by level of test anxiety with ethnicity, family composition, and class academic level, there is a strong relationship with gender. Thus, while more than half of the males (51.9%) and just over one-quarter of the females (28.2%) had a low level test anxiety, twice as many females as males had a high level of test anxiety (34.5% compared to 16.6%, respectively).

Table VI.3 Levels of Test Anxiety by Selected Characteristics

		Test Anxiety Level		
	N	Low	Med	High
All Students	745	39.3%	34.5%	26.2%
Gender				
Male	349	51.9%	31.5%	16.6%
Female	387	28.2%	37.0%	34.9%
Ethnicity				
Caucasian	314	42.0%	34.1%	23.9%
Hispanic	268	35.8%	33.6%	30.6%
Other	153	40.5%	35.9%	23.5%
Family Composition				
Both Bio Parents	470	39.8%	34.7%	25.5%
Other Parents	269	38.7%	34.2%	27.1%
Class Academic Level				
Regular	491	38.3%	33.0%	28.7%
Advanced	254	41.3%	37.4%	21.3%

Baseline Bivariate Analysis

In this section we report the results of a baseline pre-study single factor analysis of variance (ANOVA) of the thirteen primary scales constructed from the SOS conducted on the whole sample. Certain variables descriptive of the student and the academic level of their class (gender, ethnicity, family composition, and academic level) are used as the independent or grouping variable. To aid the reader, the data are presented in both tabular and graphic form for most results.

Gender

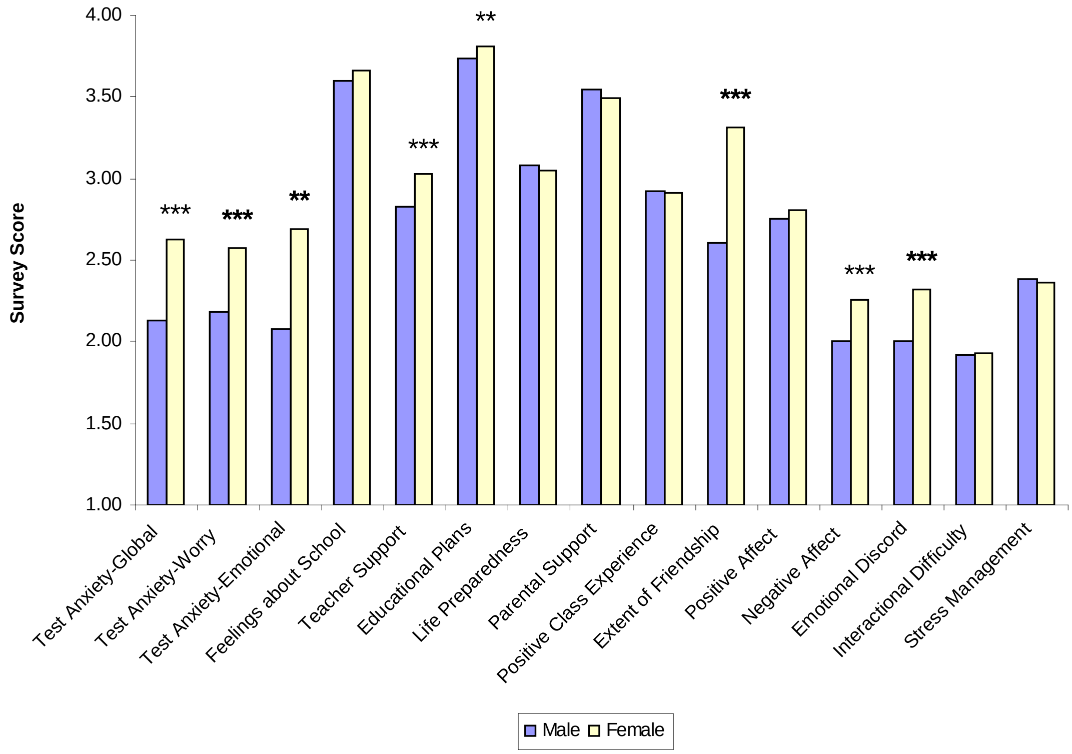
Starting with the results by gender (Table VI.4, Figure VI.4), a number of statistically significant differences are apparent in which the means for females are higher than those for males. Thus, on average, compared with their male classmates, females scored higher on their 9th grade English-Language Arts (ELA) standardized test (367.23 versus 351.39, p <0.001), reported being more anxious about taking tests (global test anxiety measure: 2.63 versus 2.13, p <0.001), felt more supported by their teacher or another adult at the school (3.03 versus 2.82, p <0.001), were more optimistic about their future education and life (3.81 versus 3.74, p <0.01), had a deeper relationship with their closest friends (3.32 versus 2.61, p <0.001), and reported higher levels of both negative affect (2.26 versus 2.00, p <0.001) and emotional discord (2.32 versus 2.00, p <0.001).

Table VI.4 Baseline Gender Comparisons

	Male (N=351)		Female (N=389)				
	Mean	SD	Mean	SD	Mean Sq.	F	*p* <
CST English - Language Arts 9	351.39	50.78	367.23	52.29	43793.80	16.45	0.001
Test Anxiety-Global	2.13	0.85	2.63	0.85	45.24	62.25	0.001
Test Anxiety-Worry	2.18	0.91	2.57	0.90	27.47	33.49	0.001
Test Anxiety-Emotional	2.08	0.91	2.69	0.92	68.32	81.65	0.001
Feelings about School	3.60	0.72	3.67	0.68	0.91	1.88	ns
Teacher Support	2.82	0.79	3.03	0.76	7.94	13.22	0.001
Educational Plans	3.74	0.43	3.81	0.35	1.06	7.08	0.01
Life Preparedness	3.08	0.71	3.05	0.69	0.17	0.36	ns
Parental Support	3.55	0.62	3.49	0.67	0.47	1.14	ns
Positive Class Experience	2.93	0.69	2.91	0.67	0.03	0.06	ns
Extent of Friendship	2.61	0.69	3.32	0.48	92.52	268.30	0.001
Positive Affect	2.75	0.61	2.80	0.62	0.52	1.36	ns
Negative Affect	2.00	0.67	2.26	0.71	11.88	24.93	0.001
Emotional Discord	2.00	0.70	2.32	0.73	18.30	35.63	0.001
Interactional Difficulty	1.92	0.61	1.93	0.61	0.05	0.14	ns
Stress Management	2.39	0.61	2.37	0.63	0.06	0.15	ns

Single Factor ANOVA

Figure VI.4 Baseline Gender Comparisons



Ethnicity

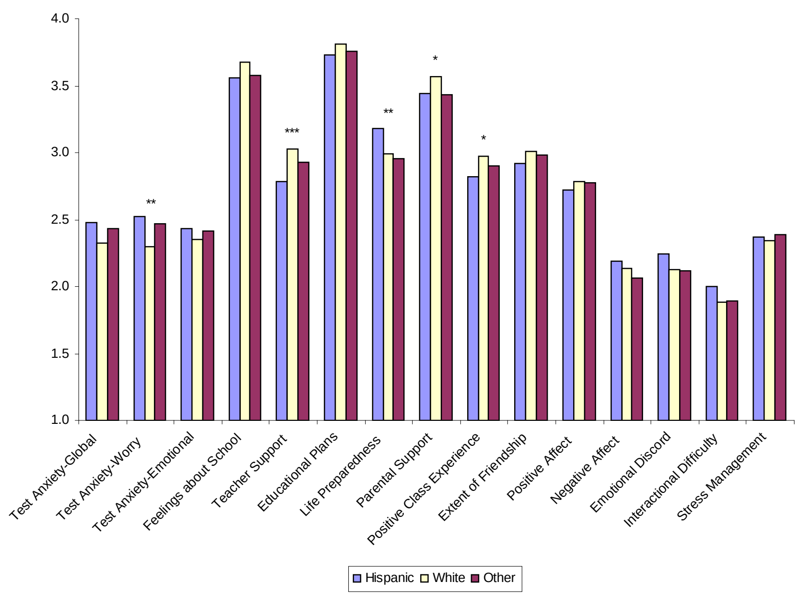
While fewer differences appear by ethnicity (Table VI.5, Figure VI.5), the significantly higher mean scores achieved by the White students over their Hispanic classmates on their 9th grade CST ELA standardized test (373.97 versus 334.70, p <0.001) and on their 10th grade CAHSEE ELA (402.84 versus 376.19, p <0.001) and CAHSEE Math (401.53 versus 373.83, p <0.001) are consistent with historical patterns in standardized test data from national sources. Hispanic students also tend to worry more about taking tests (Test Anxiety-Worry: 2.53 versus 2.30, p <0.01), are less likely to feel they have support from a teacher or an adult at school (2.79 versus 3.02, p <0.001), feel they have less support from a parent or adult at home (3.44 versus 3.57, p <0.05), and have a slightly less positive attitude about their class than their White classmates (2.82 versus 2.97, p <0.05). On the other hand, Hispanic students are slightly more likely to report feeling better prepared for life from what they learned at school than White students (3.18 versus 2.99, p <0.01).

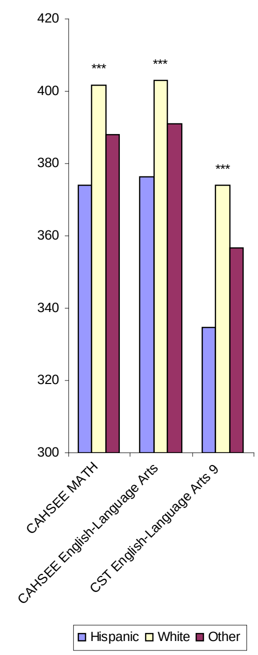
Table VI.5 Baseline Ethnicity Descriptive Statistics

	Hispanic				White				Other						
	N	Mean	SD	SEM	N	Mean	SD	SEM	N	Mean	SD	SEM	Mean Sq	F	*p* <
CAHSEE MATH	260	373.83	34.95	2.17	311	401.53	33.02	1.87	96	388.09	37.74	3.85	59804.64	51.56	0.001
CAHSEE English-Language Arts	262	376.19	31.17	1.93	311	402.84	29.28	1.66	93	390.94	33.74	3.50	48799.97	52.95	0.001
CST English-Language Arts 9	251	334.70	46.14	2.91	302	373.97	49.45	2.85	91	356.55	53.01	5.56	100217.63	43.01	0.001
Test Anxiety-Global	268	2.48	0.89	0.05	314	2.32	0.90	0.05	96	2.43	0.90	0.09	1.35	1.72	ns
Test Anxiety-Worry	268	2.53	0.93	0.06	314	2.30	0.93	0.05	96	2.47	0.92	0.09	3.98	4.72	0.01
Test Anxiety-Emotional	268	2.43	0.94	0.06	314	2.35	0.99	0.06	95	2.42	0.98	0.10	0.60	0.64	ns
Feelings about School	269	3.56	0.71	0.04	317	3.67	0.68	0.04	96	3.57	0.81	0.08	1.08	2.17	ns
Teacher Support	269	2.79	0.80	0.05	317	3.02	0.78	0.04	96	2.92	0.80	0.08	3.31	5.44	0.001
Educational Plans	268	3.73	0.43	0.03	317	3.81	0.36	0.02	96	3.75	0.42	0.04	0.34	2.22	ns
Life Preparedness	269	3.18	0.67	0.04	315	2.99	0.70	0.04	96	2.95	0.78	0.08	2.14	4.41	0.01
Parental Support	260	3.44	0.69	0.04	307	3.57	0.62	0.04	89	3.43	0.66	0.07	1.29	3.10	0.05
Positive Class Experience	259	2.82	0.68	0.04	307	2.97	0.68	0.04	89	2.90	0.72	0.08	1.51	3.26	0.05
Extent of Friendship	269	2.92	0.71	0.04	317	3.01	0.69	0.04	95	2.99	0.64	0.07	0.79	1.69	ns
Positive Affect	269	2.72	0.64	0.04	317	2.78	0.60	0.03	96	2.78	0.61	0.06	0.84	2.20	ns
Negative Affect	269	2.19	0.73	0.04	317	2.14	0.69	0.04	96	2.06	0.69	0.07	0.77	1.59	ns
Emotional Discord	269	2.24	0.74	0.05	316	2.13	0.74	0.04	95	2.12	0.71	0.07	0.77	1.43	ns
Interactional Difficulty	269	2.00	0.65	0.04	317	1.89	0.59	0.03	96	1.89	0.60	0.06	0.86	2.32	ns
Stress Management	267	2.37	0.64	0.04	317	2.34	0.61	0.03	96	2.38	0.62	0.06	0.76	1.96	ns

Single Factor ANOVA

Figure VI.5 Baseline Ethnicity Descriptive Statistics





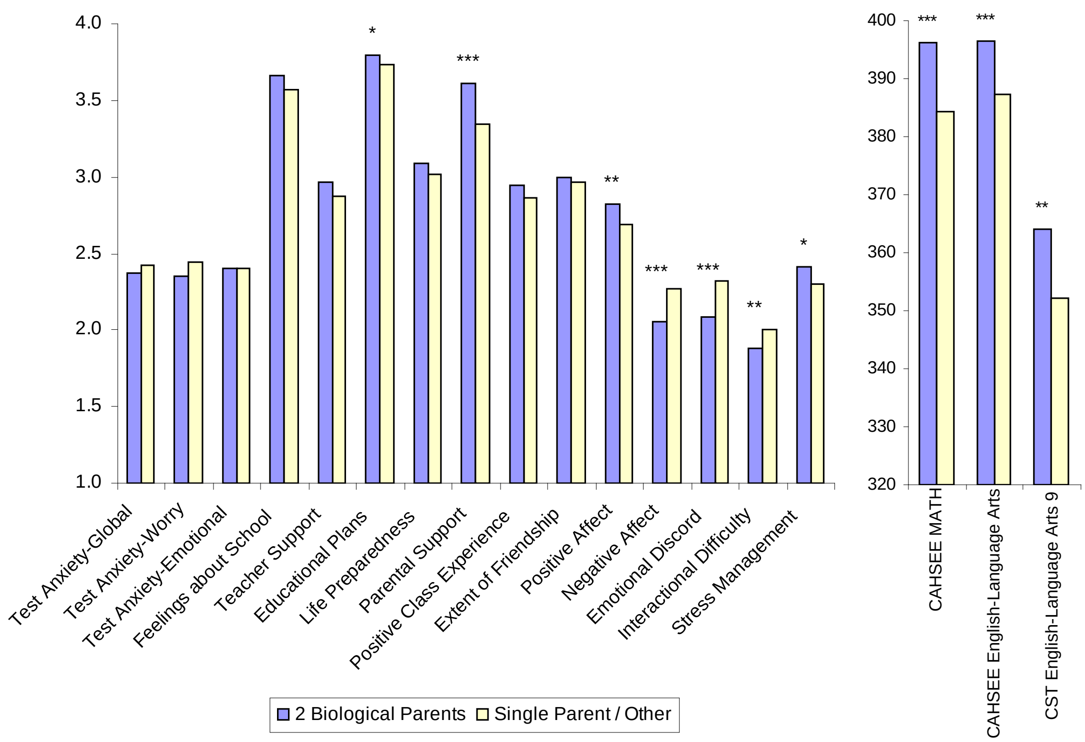
Family Composition

In terms of family composition (Table VI.6, Figure VI.6), the pattern of statistically significant differences provides evidence of a number benefits resulting from living in an intact family situation. Thus, when compared to other living arrangements, those students who live with both biological parents were more likely to be placed in an advanced academic level class (40% versus 24%, respectively), had achieved a higher 9th grade CST ELA test score (364.14 versus 352.17, $p < 0.01$) and higher CAHSEE scores in ELA (396.61 versus 387.30, $p < 0.001$) and Math (396.34 versus 384.31, $p < 0.001$), were slightly more optimistic about their future education and life (3.80 versus 3.74, $p < 0.05$), felt they had greater support from a parent or adult at home (3.61 versus 3.34, $p < 0.001$), had a slightly higher level of positive affect (2.82 versus 2.69, $p < 0.01$), and reported having a somewhat more effective stress management ability (2.42 versus 2.30, $p < 0.05$) than their classmates living in other familial situations. By contrast, the latter were more likely to report higher levels of negative affect (2.27 versus 2.05, $p < 0.001$), greater emotional discord (2.32 versus 2.08, $p < 0.001$), and greater levels of interactional difficulty (2.00 versus 1.88, $p < 0.01$) in their lives during the few weeks prior to completing the survey.

Table VI.6 Baseline Parental Situation Descriptive Statistics

	2 Biological Parents				Single Parent / Other						
	N	Mean	SD	SEM	N	Mean	SD	SEM	Mean Sq	F	*p* <
CAHSEE MATH	464	396.34	37.48	1.74	270	384.31	36.12	2.20	24681.00	18.04	0.001
CAHSEE English-Language Arts	464	396.61	33.39	1.55	269	387.30	32.74	2.00	14760.44	13.43	0.001
CST English-Language Arts 9	452	364.14	52.32	2.46	258	352.17	51.73	3.22	23522.48	8.66	0.01
Test Anxiety-Global	470	2.38	0.88	0.04	275	2.42	0.90	0.05	0.37	0.47	ns
Test Anxiety-Worry	470	2.36	0.91	0.04	275	2.45	0.95	0.06	1.42	1.67	ns
Test Anxiety-Emotional	470	2.40	0.97	0.04	274	2.41	0.95	0.06	0.01	0.01	ns
Feelings about School	473	3.66	0.69	0.03	276	3.57	0.74	0.04	1.40	2.79	ns
Teacher Support	473	2.97	0.78	0.04	276	2.87	0.80	0.05	1.48	2.40	ns
Educational Plans	472	3.80	0.37	0.02	276	3.74	0.43	0.03	0.68	4.47	0.05
Life Preparedness	472	3.08	0.69	0.03	275	3.02	0.71	0.04	0.83	1.68	ns
Parental Support	459	3.61	0.57	0.03	263	3.34	0.73	0.05	12.62	31.23	0.001
Positive Class Experience	458	2.94	0.67	0.03	263	2.87	0.72	0.04	1.07	2.28	ns
Extent of Friendship	473	3.00	0.71	0.03	275	2.96	0.65	0.04	0.21	0.45	ns
Positive Affect	473	2.82	0.64	0.03	276	2.69	0.58	0.03	3.18	8.41	0.01
Negative Affect	473	2.05	0.66	0.03	276	2.27	0.74	0.04	8.60	18.01	0.001
Emotional Discord	472	2.08	0.71	0.03	275	2.32	0.75	0.05	10.01	18.98	0.001
Interactional Difficulty	473	1.88	0.60	0.03	276	2.00	0.61	0.04	2.57	6.97	0.01
Stress Management	471	2.42	0.63	0.03	276	2.30	0.61	0.04	2.36	6.11	0.05

Figure VI.6 Baseline Parental Situation Descriptive Statistics



Class Academic Level

When the data are partitioned by class academic level (Table VI.7 and Figure VI.7), the pattern of statistically significant differences portrays contrasting realities for students. When compared to students placed in a regular class, those in an advanced class scored much higher on average on their 9th grade CST ELA test (403.82 versus 336.10, p <0.001), held more positive feelings about their school (3.79 versus 3.54 p <0.001), felt they had more support from a teacher or an adult at their school (3.13 versus 2.83, p <0.001), were more optimistic about their future education and life (3.87 versus 3.73, p <0.001), felt they had more support from a parent or adult at home (3.67 versus 3.43, p <0.001), felt more positive about their class (3.02 versus 2.86, p <0.01), reported a deeper relationship with their closest friends (3.02 versus 2.86, p <0.001), and reported higher levels of positive affect (2.88 versus 2.72, p <0.001). By contrast, those in regular classes were more likely to worry about taking tests (Test Anxiety-Worry: 2.49 versus 2.20, p <0.001), reported higher levels of negative affect (2.17 versus

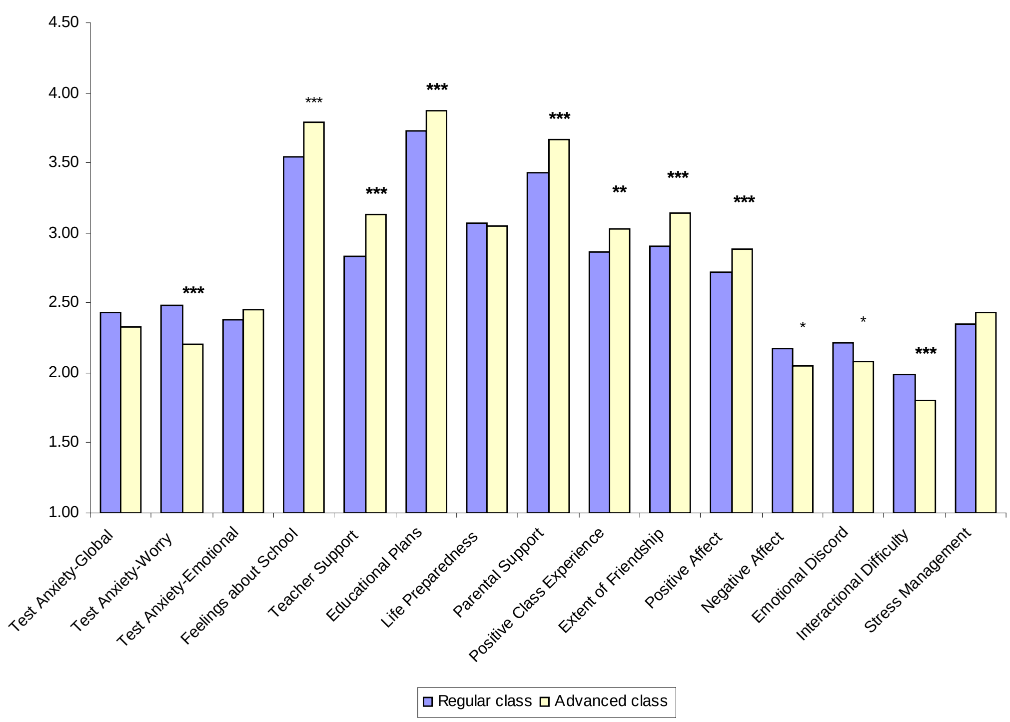
2.05, $p < 0.05$), felt greater emotional discord (2.22 versus 2.08, $p < 0.05$), and reported having greater difficulty in their interactions with others (1.99 versus 1.81, $p < 0.001$) in the weeks prior to completing the survey.

Table VI.7 Baseline Test Performance, Test Anxiety, and SOS Scales by Class Academic Level

	Regular Students (N=494)			Advanced Students (N=255)					
	Mean	SD	SEM	Mean	SD	SEM	Mean Square Btw. Grp	F	$p <$
CST English - Language Arts 9	336.10	43.22	2.01	403.82	37.08	2.35	741435.66	437.35	0.001
Test Anxiety-Global	2.43	0.91	0.04	2.32	0.84	0.05	1.92	2.45	ns
Test Anxiety-Worry	2.49	0.94	0.04	2.20	0.86	0.05	13.64	16.28	0.001
Test Anxiety-Emotional	2.38	0.97	0.04	2.45	0.95	0.06	0.79	0.85	ns
Feelings about School	3.54	0.71	0.03	3.79	0.68	0.04	10.29	21.02	0.001
Teacher Support	2.83	0.79	0.04	3.13	0.74	0.05	15.55	25.97	0.001
Educational Plans	3.73	0.43	0.02	3.87	0.27	0.02	3.35	22.52	0.001
Life Preparedness	3.07	0.71	0.03	3.04	0.70	0.04	0.09	0.17	ns
Parental Support	3.43	0.70	0.03	3.67	0.51	0.03	9.23	22.57	0.001
Positive Class Experience	2.86	0.70	0.03	3.02	0.65	0.04	4.49	9.68	0.01
Extent of Friendship	2.90	0.71	0.03	3.14	0.61	0.04	9.42	20.59	0.001
Positive Affect	2.72	0.63	0.03	2.88	0.59	0.04	4.54	12.07	0.001
Negative Affect	2.17	0.73	0.03	2.05	0.62	0.04	2.38	4.89	0.05
Emotional Discord	2.22	0.74	0.03	2.08	0.72	0.04	3.12	5.81	0.05
Interactional Difficulty	1.99	0.64	0.03	1.81	0.52	0.03	5.62	15.39	0.001
Stress Management	2.35	0.63	0.03	2.43	0.61	0.04	1.00	2.59	ns

Single Factor ANOVA

Figure VI.7 Baseline Test Performance, Test Anxiety, and SOS Scales by Class Academic Level



Operational Problems in Using Standardized Tests

Before examining the relationship between test anxiety and test performance, it is necessary to mention a number of operational problems we encountered in our efforts to measure student test performance with the use of standardized test scores. Students at both high schools take two standardized tests: the California High School Exit Examination (CAHSEE)—a proficiency test, taken midway through the term; and the California Standardized Test (CST)—an achievement test, taken at the end of the term. Because the CAHSEE was administered well before teachers had completed training students in the TestEdge intervention, we had planned to use the CST as our measure of student test performance—the dependent variable for the study. The plan was to use 9th grade CST scores to establish a baseline measure of each student's pre-study test performance level and compare them with the 10th grade CST scores to measure the degree of improvement in performance. However, with the exception of the CST English-Language Arts test, which appeared to be administered universally on a standardized basis to all students in both 9th and 10th grades, and thus met our need for a repeated measures format, a number of unanticipated complications prevented our use of much of the CST data.

The main problem was the inability to construct a standardized, repeated measures framework from the highly variable combination of tests in different CST subject areas that students had taken. Not only were different combinations of subject areas taken in the experimental group than in the control group, but many students also took different combinations of subject area tests in the 9th and 10th grades. For example, in the 9th grade 91% of the experimental group took Earth Science while 85% of the control group took Biology; in the 10th grade most of the experimental group took Biology while the control group took Chemistry. This meant that the CST Science scores could not be compared and thus were unusable. While these difficulties were rife in the CST in Mathematics (see Table VI.8), in that different groups of students took various combinations of General Math, Summative Math, Algebra 1, Geometry, and Algebra 2, we found a notable subset of 183 students (121 in the experimental group and 62 in the control group) who all took Geometry in the 9th grade and who also all took Algebra 2 in the 10th grade. In the analyses that follow in a later section, this group of students is referred to as Math Group 1.

Table VI.8 CST Math Test Type

	Experimental		Control	
	Math 9	Math 10	Math 9	Math 10
General math	25%	0%	25%	0%
Summative math	0%	0%	0%	17%
Algebra 1	43%	29%	28%	15%
Geometry	28%	40%	34%	27%
Algebra 2	0%	26%	13%	40%
Unknown	3%	5%	0%	0%

Relationship Between Test Anxiety and Test Performance

A major focus of the study is the relationship between test anxiety and test performance. However, while there is a consistent body of evidence from previous research of an association between test anxiety and test performance, there is as yet no definitive evidence of a direct causal relationship between the two (Spielberger & Vagg, 1995b; Zeidner, 1998). In this section, we present evidence confirming the negative relationship between test anxiety and test performance. In presenting the results of the physiological study, in Chapter XI, below, we will examine the evidence for a causal relationship between electrophysiological measures of test stress and test performance.

In order to validate the use of our abbreviated eight-item version of the sixteen-item Spielberger Test Anxiety Inventory, it is not only necessary to empirically document that it achieves a comparable level of measurement validity and reliability as the original instrument, as was shown above, but also to show that it has the same expected relationship with performance on standardized tests. This is shown in Table VI.9 and Figure VI.9, where students' test anxiety (global), CAHSEE–Mathematics, and CAHSEE–English-Language Arts scores have each been classified into three approximately equal-size groupings of low, medium, and high scores. A strong, statistically significant ($p < 0.001$), negative relationship is clearly apparent between test anxiety level and level of test performance on the two CAHSEE tests. For example, students with the lowest test anxiety (mean TAI = 1.42) achieved the highest scores on both Math (mean = 400.73) and ELA (mean = 400.93); by contrast students with the highest test anxiety (mean TAI = 3.43) produced the lowest test scores on Math (mean = 384.13) and ELA (386.62). In other words, there is a 15-point difference on both tests between students in the high and low test anxiety categories.

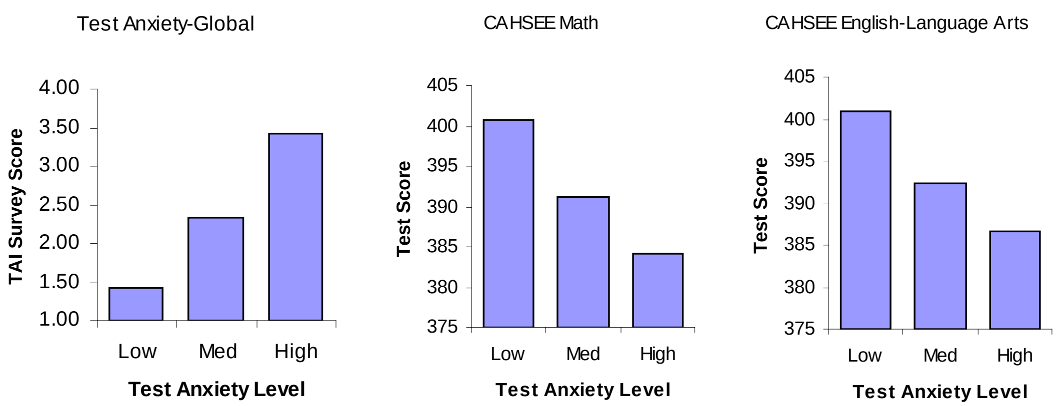
Table VI.9 Graduate Exam Scores by Baseline Test Anxiety Level

	Low TAI (N=256)			Medium TAI (N=234)			High TAI (N=255)				
	Mean	SEM	SD	Mean	SEM	SD	Mean	SEM	SD	F	*p* <
Test Anxiety-Global	1.42	0.02	0.28	2.34	0.02	0.24	3.43	0.02	0.37	2782.75	0.001
CAHSEE Math	400.73	2.33	36.91	391.19	2.41	36.47	384.13	2.33	36.88	12.86	0.001
CAHSEE English-Language Arts	400.90	2.14	33.77	392.30	2.10	31.89	386.62	2.08	32.95	11.88	0.001

Single factor ANOVA

Bonferroni multiple comparisons. All pairwise comparisons were significant except the difference between the Med & High Math score pair and the Med & High ELA score pair.

Figure VI.9 Graduate Exam Scores by Baseline Test Anxiety Level



Additional evidence of this inverse relationship between test anxiety and test performance in the whole sample is presented in Table VI.10. Two distinct sets of Pearson product moment correlations are shown, one showing the relationship between student test anxiety and test performance on each of the four CAHSEE and CST performance measures (bottom left quadrant of the table), and another showing the intercorrelations between each pair of tests (bottom right triangle). For the former, a consistent pattern of statistically significant negative correlations (all correlation coefficients are in the 0.2 range) is apparent between the global measure of test anxiety and student performance on each test. Consistent with prior research, it is worth noting that the "worry" component of test anxiety appears to be that which has the strongest inverse relationship to test performance (correlation coefficients range from -0.288 to -0.340).

Turning to the intercorrelations among all pair-wise combinations of the four tests, a strong, consistent pattern of statistically significant high correlations is evident for each pairing, which range from a low of 0.621 between 9th grade CST ELA score and 10th grade CST Math score, to a high of 0.846 between 9th grade CST ELA score and 10th grade CST ELA score. In short, student performance on any one of these standardized tests is highly related to their performance on any other standardized test—a finding entirely consistent with what is known about student test performance.

Table VI.10 Baseline Test Anxiety and Performance Correlations (N=745)

	1	2	3	4	5	6	7	8	9
1. Test Anxiety-Global	1	NC	NC	-.200 **	-.199 **	-.240 **	-.205 **	-.248 **	-.231 **
2. Test Anxiety-Worry	NC	1	NC	-.296 **	-.288 **	-.322 **	-.288 **	-.340 **	-.330 **
3. Test Anxiety-Emotional	NC	NC	1	-.083 *	-.089 *	-.133 **	-.101 **	-.130 **	-.107 **
4. Math 9 (CAHSEE)	-.200 **	-.296 **	-.083 *	1	.774 **	.760 **	.743 **	.725 **	.746 **
5. English-Language Arts (CAHSEE)	-.199 **	-.288 **	-.089 *	.774 **	1	.650 **	.641 **	.816 **	.816 **
6. CST Math 9th grade	-.240 **	-.322 **	-.133 **	.760 **	.650 **	1	.703 **	.665 **	.658 **
7. CST Math 10th grade	-.205 **	-.288 **	-.101 **	.743 **	.641 **	.703 **	1	.621 **	.627 **
8. CST English-Language Arts 9th grade	-.248 **	-.340 **	-.130 **	.725 **	.816 **	.665 **	.621 **	1	.846 **
9. CST English-Language Arts 10th grade	-.231 **	-.330 **	-.107 **	.746 **	.816 **	.658 **	.627 **	.846 **	1

** $p < 0.01$ level (2-tailed).

* $p < 0.05$ level (2-tailed).

NC = Not computed

Explaining Test Anxiety and Test Performance

To complete the analysis of the whole sample, a series of stepwise multiple regression analyses were conducted to build an initial empirical understanding of the degree to which test anxiety and test performance were explained by three of the SOS sociodemographic variables (Gender, Ethnicity, and Family Composition) and the SOS scales. Two sets of analyses were conducted: one in which test anxiety was treated as the dependent variable, and another in which test performance (both on the CAHSEE and CST) was the dependent variable.

Starting with test anxiety, two stepwise multiple regression analyses (results not shown) were conducted using different measures of test anxiety. For the first analysis, the degree of change in test anxiety from the baseline pre-study measurement point was used as the dependent variable. Only one factor met the criteria for entry into the statistical model—Emotional Discord—and it only explained about 1% of the change in test anxiety (adjusted R^2 = 0.011, standardized beta coefficient = 0.114). For the second analysis, the degree of change in test anxiety between the pre and post study moments was used as the dependent variable. This time, two factors were included in the statistical model—Emotional Discord and Interactional Difficulty. While there was some improvement in explanatory power, together they still only explained approximately 5% of the measured change in test anxiety (adjusted $R^2 = 0.054$, standardized beta coefficients = 0.195 and 0.089, respectively).

Turning to test performance, for these analyses, the "Worry" and "Emotionality" test anxiety sub-scales were used in place of the TAI-Global total scale, both because the Worry component was found to be more highly correlated with test performance than the global measure, and also because another series of regression analyses (not shown) had found that TAI-Global possessed less explanatory power than the "worry" subscale we report here.

Beginning with the results for the CAHSEE (Table VI.11), with the exception of Gender and Negative Affect included in the regression model for the ELA test, both the ELA and Math tests have regression models that share seven factors in common: Feelings About School, Test Anxiety-Worry, Test Anxiety-Emotionality, Life Preparedness, Educational Plans, Teacher Support, and Family Composition. The nine factors in the regression model for ELA explain approximately 24% of the variance in student test performance (adjusted $R^2 = 0.24$); the seven factors in the model for Math explain approximately 20% of the variance in student test performance (adjusted $R^2 = 0.20$).

Five of these seven factors also appear in the pre-study, post-study, and mid-study regression models for the CST (Table VI.12): Test Anxiety-Worry, Feelings about School, Test Anxiety-Emotionality, Life Preparedness, and Educational Plans. Also, Gender is a common factor in the pre-study, mid-study, and post-study regression models for the CST. Each of the three regression models explains almost one-quarter of the variance in student test performance. Thus, in the pre-study model of the 9th grade ELA test, the adjusted R^2 was 0.24; in the post-study model of 10th grade ELA the adjusted R^2 was 0.23; and in the mid-study model of 10th grade ELA it was 0.24.

It is worth noting that a further set of regression analyses (not shown) found that, by itself, Test Anxiety-Worry accounted for about one third to almost one half of the variance in 9th grade test performance on the two tests explained by the pre-study regression models; for the 9th grade CAHSEE Math and CAHSEE ELA the adjusted R^2 was 0.076 and 0.074, respectively (standardized Beta coefficients were -0.278 and -0.2740), and for the 9th grade CST Math and CST ELA the adjusted R^2 was 0.0976 and 0.103, respectively (standardized Beta coefficients were -0.323 and -0.311).

Table VI.11.a Stepwise Regression of CAHSEE ELA Test Scores

Dependent Variable: CAHSEE English-Language Arts
Stepwise Regression Model Summary

	Unstandardized Coefficients		Standardized	t	*p*
	B	Std. Error	Beta		
(Constant)	325.39	11.79		27.60	0.000
Feelings about School	9.29	1.83	0.21	5.07	0.000
Test Anxiety-Worry	-16.13	1.97	-0.43	-8.19	0.000
Gender	11.18	2.41	0.17	4.64	0.000
Test Anxiety-Emotional	8.05	1.88	0.23	4.27	0.000
Life Preparedness	-12.22	2.01	-0.25	-6.07	0.000
Educational Plans	13.03	2.83	0.18	4.61	0.000
Teacher Support	5.51	1.70	0.13	3.23	0.001
Negative Affect	4.55	1.89	0.10	2.41	0.016
Both Bio Parents	-1.43	0.67	-0.07	-2.14	0.033

R	R Square	Adj. R Square	Std. Error of the Estimate
0.50	0.25	0.24	29.28

Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).

Table VI.11.b Stepwise Regression of CAHSEE Math Test Scores

Dependent Variable: CAHSEE Math
Stepwise Regression Model Summary

	Unstandardized Coefficients		Standardized	t	*p*
	B	Std. Error	Beta		
(Constant)	354.91	12.12		29.29	0.000
Feelings about School	9.50	2.04	0.20	4.65	0.000
Test Anxiety-Worry	-18.58	2.24	-0.44	-8.28	0.000
Test Anxiety-Emotional	12.52	2.09	0.32	5.98	0.000
Both Bio Parents	-2.50	0.76	-0.11	-3.27	0.001
Life Preparedness	-10.96	2.28	-0.20	-4.80	0.000
Teacher Support	6.00	1.94	0.13	3.10	0.002
Educational Plans	9.83	3.23	0.12	3.04	0.002

R	R Square	Adj. R Square	Std. Error of the Estimate
0.45	0.20	0.20	33.54

Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).

Table VI.12 Pre-Study, Post-Study, and Mid-Study Regression Models for CST ELA Test Scores

Table VI.12.a Pre-Study

Dependent Variable: CST English - Language Arts 9
Stepwise Regression Model Summary

	Unstandardized Coefficients		Standardized	t	*p*
	B	Std. Error	Beta		
(Constant)	277.71	19.95		13.92	0.000
Test Anxiety-Worry	-29.22	3.02	-0.51	-9.69	0.000
Gender	17.83	3.77	0.17	4.73	0.000
Feelings about School	13.66	3.09	0.18	4.42	0.000
Test Anxiety-Emotional	11.97	2.97	0.22	4.03	0.000
Life Preparedness	-16.73	3.16	-0.22	-5.29	0.000
Positive Class Experience	10.30	2.97	0.13	3.47	0.001
Educational Plans	18.01	5.29	0.13	3.40	0.001

R	R Square	Adj. R Square	Std. Error of the Estimate
0.50	0.25	0.24	45.65

Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).

Table VI.12.b Post-Study

Dependent Variable: CST English - Language Arts 10
Stepwise Regression Model Summary

	Unstandardized Coefficients		Standardized	t	*p*
	B	Std. Error	Beta		
(Constant)	273.76	18.29		14.97	0.000
Test Anxiety-Worry	-28.10	3.28	-0.45	-8.56	0.000
Test Anxiety-Emotional	16.05	3.08	0.27	5.22	0.000
Gender	16.82	3.95	0.15	4.26	0.000
Educational Plans	15.69	4.93	0.13	3.18	0.002
Life Preparedness	-18.56	3.39	-0.23	-5.47	0.000
Feelings about School	11.12	2.98	0.15	3.73	0.000
Teacher Support	8.66	2.85	0.12	3.04	0.002
Parental Support	9.92	3.35	0.12	2.96	0.003
Stress Management	-8.04	3.23	-0.09	-2.49	0.013

R	R Square	Adj. R Square	Std. Error of the Estimate
0.49	0.24	0.23	48.67

Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).

Table VI.12.c Mid-Study

Dependent Variable: CST English - Language Arts 10
Stepwise Regression Model Summary

	Unstandardized Coefficients		Standardized	t	*p*
	B	Std. Error	Beta		
(Constant)	250.02	22.36		11.18	0.000
Test Anxiety-Worry (Time 3)	-35.08	3.97	-0.48	-8.84	0.000
Gender	16.96	4.13	0.16	4.10	0.000
Test Anxiety-Emotional (Time 3)	15.80	3.46	0.24	4.57	0.000
Parental Support	10.45	3.43	0.12	3.05	0.002
Life Preparedness	-20.82	3.44	-0.27	-6.05	0.000
Teacher Support	9.49	2.93	0.14	3.24	0.001
Educational Plans	18.18	5.20	0.14	3.49	0.001
Feelings about School	9.90	3.12	0.14	3.17	0.002
Negative Affect	9.04	3.19	0.12	2.84	0.005
Q4. Intact family	-2.76	1.21	-0.08	-2.29	0.022

	R	R Square	Adj. R Square	Std. Error of the Estimate
	0.50	0.25	0.24	47.39

Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).

Finally, we conducted a regression analysis (results not shown) with the degree of change in student test performance on CST ELA from the 9th grade to the 10th grade as the dependent variable. Using the stepwise procedure, the three sociodemographic variables and all thirteen SOS scales were considered for entry into the statistical model. With one exception, none of the other factors met the entry criteria (p of F ≤ 0.05; p of F to remove ≥ 0.10). This exception was Positive Affect, which explained little of the observed change in student test scores (adjusted $R^2 = 0.005$, standardized beta coefficient $= 0.083$).

Affective Mood, Social Behavior, Test Anxiety, and Test Performance

An independent analysis of the relationship between the SOS measures of affective mood, social behavior, and test anxiety and student test performance was conducted by Kimberly Hartnett-Edwards (2006), another CGU graduate student member of the TENDS research team. Instead of using the thirteen multivariate scales (Life Preparedness, Teacher Support, Emotional Discord, etc.) we constructed from the SOS variables, she grouped the individual items into two alternate constructs—Affective Mood and Social Behavior—and also used the individual items in the Test Anxiety Inventory. She then conducted a series of stepwise multiple regression analyses to predict 10th grade test performance on both the CST and CAHSEE English-Language Arts exams, using the individual items in each of the three constructs as independent variables (see Hartnett-Edwards (2006: 120-135). She found that the measures of emotional disposition in her Affective Mood construct explained *twice* the amount of variance in test performance as the items in the test anxiety scale.

Starting with her results for Affective Mood (Hartnett-Edwards, 2006, Tables 23 and 24), ten and eleven variables, respectively, met the stepwise inclusion criteria for entry into the two regression models. The model for test performance on the CST-ELA explained approximately 20% of the variance ($R^2 = 0.202$; F = 19.927, p <0.000) and the model for the CAHSEE-ELA explained approximately 23% of the variance ($R^2 = 0.226$; F = 20.896, p <0.000). Seven variables were common to both models (the standardized beta coefficient for the CST model and the CAHSEE model, respectively, are in parentheses):

I plan to go to college or some other school after high school (0.229, 0.214)

I feel that there are mostly good feelings among all of us in this class (0.171, 0.15)

I feel what I have learned at this school has inspired me to want more for myself out of life (-0.218, -0.191)

I am happy to be at this school (0.134, 0.172)

In my home, there is a parent or some other adult who is interested in my schoolwork (0.07, 0.102)

I feel stressed (0.15, 0.146)

I feel angry (-0.096, -0.122)

Moving on to her results for Social Behavior (Hartnett-Edwards, 2006, Tables 24 and 25), four variables common to both regression models met the stepwise inclusion criteria and explained approximately 11–12% of the variance in student test performance (CST-ELA model: $R^2 = 0.122$; F = 27.574, *p* <0.000; CAHSEE-ELA model: $R^2 = 0.113$; F = 25.217, *p* <0.000):

For those classmates or friends I feel closest to, we have strong feelings about each other (0.24, 0.179)

In my interactions with others, I get into arguments or fights (-0.172, -0.186)

For those classmates or friends I feel closest to, I talk on the phone (-0.194, -0.159)

For those classmates or friends I feel closest to, we feel very comfortable with each other (0.124, 0.157)

Finally, her results for Test Anxiety (Hartnett-Edwards, 2006, Tables 28 and 29) showed that five of the individual items in the TAI, common to both regression models, met the stepwise inclusion criteria and explained approximately 14% and 12%, respectively, of the variance in student test performance (CST-ELA model: $R^2 = 0.142$; $F = 26.266$, $p < 0.000$; CAHSEE-ELA model: $R^2 = 0.115$; $F = 20.533$, $p < 0.000$):

I freeze up on important exams (-0.253, -0.239)

The harder I work at a test, the more confused I get (-0.148, -0.149)

During tests, I feel very tense (0.171, 0.106)

During tests, I find myself thinking about the consequences of failing (-0.199, -0.138)

I worry a great deal before taking an important examination (0.146, 0.176)

What is noteworthy about Hartnett-Edwards' results is that for both tests the measures of the Affective Mood construct explain virtually twice as much of the variance in student test performance as the measures from the test anxiety scale (25% versus ~13%, respectively). As she points out, this suggests that students' overall emotional disposition—their emotional awareness and skill to appropriately manage emotions and feelings in their lives as a whole—appears to be a stronger predictor of test performance than their specific fears and worries about taking an important test. Moreover, the interactional measures of the Social Behavior construct have about the same predictive power as the measures in the test anxiety scale. This suggests also that the nature and quality of students' relationships, both in and outside the classroom, are equally important in affecting test performance as is anxiety about a particular test.

A further point Hartnett-Edwards makes about her results is that an examination of the sign and strength of the beta coefficients suggests that there is a tendency for positive feelings and emotions (e.g., mostly good feelings, happy to be at this school) and positive interactions and relations (e.g., strong feelings about each other, feel very comfortable with each other) to have a positive effect on test performance, while strongly negative feelings (e.g., I feel angry) and strongly negative interactions (e.g., I get into arguments or fights) have a negative effect on test performance.

Overall, Hartnett-Edwards' results not only corroborate the main findings from our analysis, but they also extend the understanding of the factors that explain test performance. These findings are consistent with the theoretical expectations described in

Chapter II, concerning the importance of student emotions and classroom socioemotional dynamics in affecting learning and academic performance.

Summary of Findings

- Students were most positive about their Educational Plans, Feelings about School, and Parental Support
- There is evidence of frequent experience of test anxiety for a large proportion of students; a similar picture is evident for the experience of Negative Affect and Emotional Discord
- Compared to male students, females performed better on the 9th grade ELA standardized test, experienced greater test anxiety, reported more support from their teacher or another adult at the school, were more optimistic about their future education and life, had deeper close friendships, and reported higher levels of negative affect and emotional discord
- Hispanic students scored lower on all standardized tests used in the study, worried more about taking tests, felt less likely to have support from a teacher or an adult at school or from a parent or adult at home, and had a less positive attitude about class than their White classmates
- In relation to those in other family circumstances, students from intact families were more likely to be placed in an advanced academic level class, had higher test performance on all three California tests (*viz.*, CAHSEE–ELA, CAHSEE–Math, and CST–ELA), were more optimistic about their future, felt they had support from a parent or adult at home, had a higher level of positive affect, and reported more effective stress management; by contrast, those in other family situations had higher levels of negative affect, greater emotional discord, and greater levels of interactional difficulty than students in intact families
- Compared to students in a regular class, those in an advanced academic class scored much higher on all three California tests, were more positive about their school, felt more supported by a teacher or school adult, were more optimistic about their future, felt they had greater support at home, felt more positive about their class, had closer friendships, and reported higher levels of positive affect. By contrast, students in a regular classes worried more about taking tests,

and they also reported higher levels of negative affect, emotional discord, and interactional difficulty

- A strong inverse relationship was found between test anxiety level and level of test performance on the two CAHSEE tests; a consistent pattern of negative correlations was found between test anxiety and student performance on each of the California tests
- Five common factors were found in the regression analysis models to explain student test performance on the CAHSEE and the CST: Test Anxiety-Worry, Feelings about School, Test Anxiety-Emotionality, Life Preparedness, and Educational Plans; for both tests the regression models explained about 20–24% of the variance in student test performance.
- Multiple regression analysis found that measures of Affective Mood explained approximately twice the variance in student test performance on both the CST-ELA and CAHSEE-ELA as items from the TAI (23% versus ~13%, respectively). Moreover, measures of student Social Behavior have about the same predictive power (~12%) as measures from the test anxiety scale. Positive feelings and emotions and prosocial behaviors have a positive effect on test performance, while strongly negative feelings and antisocial behavior and interactions have a negative impact.

Chapter VII
Commonalities and Differences Within the Experimental and Control Schools

While, in terms of the study's primary objective, evidence of a successful intervention in the experimental group requires evidence of a reduction in test anxiety covarying with an increase in test performance when compared to the control group, it is important to understand the nature and magnitude of pre-post changes observed within each school system in itself. Not only will such a within-groups analysis provide a clearer understanding of the unique endogenous conditions and forces at work in each of these two social systems, but any within-groups differences may also help explain the successes and failures of the HeartMath intervention in the experimental group.

Analysis Strategy

The analysis strategy was to start with pre and post factor analyses of the scales that constitute our primary constructs and investigate the degree to which the latent structure of the relations among these variables has changed as a result of the TestEdge intervention. While there is some evidence of similarity in the factor composition in each group, there also is evidence of two notable pre-post changes that differentiate the two groups. One involves a change in the salience of test anxiety and the second involves a change in the salience of school-related perceptions, feelings, and relations.

We followed the factor analysis with paired t-test comparisons and an ANCOVA to examine the relationship between pre-post changes in test anxiety, test performance, the various SOS scales, and the measures of sociodemographic background and class academic level. There is evidence of more extensive pre-post changes on these measures in the experimental group than were observed in the control group.

Factor Analysis and Principal Components

Results for the Experimental Group

Table VII.1 presents the results of the factor analysis for the pre- and post-study measurement. At the time of baseline data collection, the analysis identified three principal components.

The *first factor* (Component 1 in the table) was comprised of five items which are a mixture of emotions and relations:

Interactional Difficulty
Negative Affect
Emotional Discord
Positive Affect
Parental Support

It is notable that both Positive and Negative Affect are included, along with Interactional Difficulty and Parental Support. When taken as a whole, the common underlying theme in the items comprising this factor appears to be emotions and relations that are not specifically related to school.

The three measurements of test anxiety emerge as the *second factor* (Component 2):

Test Anxiety-Global
Test Anxiety-Emotionality
Test Anxiety-Worry

The *third factor* (Component 3) is comprised of the remaining seven items and, with one exception—Stress Management—appears focused around school:

Positive Class Experience
Teacher Support

Life Preparedness

Feelings about School

Extent of Friendship

Stress Management

Educational Plans

Table VII.1 Factor Analysis (with Vari-Max Rotation) of Student Opinion Survey Scales, Pre and Post-Intervention, for the Experimental Group

Experimental Group Pre
Rotated Component Matrix

	Component		
	1	**2**	**3**
Interactional Difficulty	**0.81**	0.14	-0.07
Negative Affect	**0.79**	0.31	-0.06
Emotional Discord	**0.75**	0.31	0.04
Positive Affect	**-0.65**	-0.01	0.51
Parental Support	**-0.59**	0.05	0.26
Test Anxiety-Global	0.14	**0.98**	0.02
Test Anxiety-Emotional	0.14	**0.93**	0.06
Test Anxiety-Worry	0.12	**0.92**	-0.03
Positive Class Experience	-0.01	-0.10	**0.73**
Teacher Support	-0.13	0.07	**0.69**
Life Preparedness	-0.20	-0.05	**0.68**
Feelings about School	-0.19	-0.11	**0.65**
Extent of Friendship	0.06	0.20	**0.51**
Stress Management	-0.43	0.09	**0.50**
Educational Plans	-0.44	0.07	**0.44**

Experimental Group Post
Rotated Component Matrix

	Component		
	1	**2**	**3**
Teacher Support	**0.75**	-0.14	0.01
Positive Class Experience	**0.71**	-0.06	-0.15
Life Preparedness	**0.66**	-0.26	-0.03
Feelings about School	**0.66**	-0.25	-0.07
Educational Plans	**0.61**	-0.20	-0.02
Extent of Friendship	**0.58**	0.11	0.13
Negative Affect	-0.08	**0.84**	0.26
Emotional Discord	0.08	**0.79**	0.28
Interactional Difficulty	-0.20	**0.78**	0.12
Positive Affect	0.48	**-0.62**	0.01
Stress Management	0.38	**-0.52**	0.14
Parental Support	0.41	**-0.45**	-0.05
Test Anxiety-Global	-0.02	0.14	**0.98**
Test Anxiety-Worry	-0.05	0.13	**0.93**
Test Anxiety-Emotional	0.00	0.14	**0.92**

By the end of the study, an interesting change in student perceptions and opinions is evident in the composition and structure of the principal components. While essentially the same three factors are evident, there has been a change in their ordering: items involving school have emerged as the strongest in terms of statistical power; items involving emotions and relations not directly related to school are now second; the three test anxiety measures have now moved to the bottom.

In more specific terms, with the exception of Stress Management, the *first factor* (Component 1) is now comprised of the same items as Factor Three in the pre-study analysis:

Teacher Support
Positive Class Experience
Life Preparedness
Feelings about School
Educational Plans
Extent of Friendship

The *second factor* (Component 2) has expanded from the five items included in the first factor in the pre-study analysis to six items, with the addition of Stress Management:

Negative Affect
Emotional Discord
Interactional Difficulty
Positive Affect
Stress Management
Parental Support

And finally, as already noted, test anxiety is the *third factor*:

Test Anxiety-Global
Test Anxiety-Worry
Test Anxiety-Emotional

Results for the Control Group

Table VII.2 presents the results of the factor analysis, with vari-max component rotation, for the pre and post study measurements, respectively. At the time of baseline data collection, the analysis identified three principal components.

Table VII.2 Factor Analysis (with Vari-Max Rotation) of Student Opinion Survey Scales, Pre and Post-Intervention, for the Control Group

Control Group Pre
Rotated Component Matrix

	Component		
	1	2	3
Life Preparedness	**0.71**	-0.09	-0.16
Feelings about School	**0.70**	-0.04	-0.23
Teacher Support	**0.69**	0.01	-0.07
Positive Class Experience	**0.64**	-0.16	0.03
Educational Plans	**0.61**	0.15	-0.12
Positive Affect	**0.60**	-0.08	-0.49
Parental Support	**0.54**	-0.08	-0.24
Extent of Friendship	**0.51**	0.30	0.02
Test Anxiety-Global	-0.03	**0.98**	0.17
Test Anxiety-Worry	-0.05	**0.92**	0.10
Test Anxiety-Emotional	0.00	**0.90**	0.21
Negative Affect	-0.14	0.17	**0.85**
Emotional Discord	-0.02	0.23	**0.81**
Interactional Difficulty	-0.16	0.07	**0.79**
Stress Management	0.41	-0.02	**-0.57**

Control Group Post
Rotated Component Matrix

	Component			
	1	2	3	4
Test Anxiety-Global	**0.97**	-0.12	-0.05	-0.09
Test Anxiety-Emotional	**0.91**	-0.16	-0.02	-0.06
Test Anxiety-Worry	**0.91**	-0.06	-0.07	-0.11
Stress Management	0.10	**0.78**	0.37	-0.09
Negative Affect	0.33	**-0.71**	0.01	-0.25
Positive Affect	-0.07	**0.70**	0.33	0.28
Emotional Discord	0.45	**-0.57**	0.13	-0.37
Positive Class Experience	-0.08	0.05	**0.75**	-0.11
Teacher Support	-0.09	-0.05	**0.73**	0.32
Life Preparedness	0.01	0.38	**0.63**	0.16
Feelings about School	-0.07	0.21	**0.54**	0.37
Extent of Friendship	0.34	0.16	**0.48**	0.23
Educational Plans	-0.03	0.03	0.19	**0.80**
Parental Support	-0.07	0.24	0.26	**0.62**
Interactional Difficulty	0.31	-0.43	0.02	**-0.58**

The *first and strongest factor* (Component 1 in the table), in terms of statistical power, was comprised of the following SOS scales, which are listed in order of the strength of correlation with the latent factor component:

Life Preparedness

Feelings about School

Positive Class Experience

Educational Plans

Positive Affect

Parental Support

Extent of Friendship

What appears common through these items is that they all are positive feelings and emotions and sources of social support for the student's schooling and education.

The *second factor* (Component 2) consists of the three measures of test anxiety:

Test Anxiety-Global

Test Anxiety-Worry

Test Anxiety-Emotional

Clearly, the common trait connecting the three is test anxiety.

The *third factor* (Component 3) is comprised of the remaining four items:

Negative Affect

Emotional Discord

Interactional Difficulty

Stress Management

Three of these items are negative, involving negative emotions or feelings, and problems and tensions in relationships. The fourth item, Stress Management, can be interpreted as the means of dealing with these issues and problems in the student's life.

By the end of the study—that is, at the time of post-study measurement—the results of the analysis reveal that two notable changes have occurred in the latent structure of students' perceptions. The first is that there are now four principal components, and the second is that there has been a notable rearrangement in the structure of the components.

The three measures of test anxiety have now become the *first factor* or principal component.

The *second factor* (Component 2 in the table) is similar to the third component identified in the pre-study analysis with a small but notable change in which Interactional Difficulty has now been replaced with Positive Affect:

Stress Management
Negative Affect
Positive Affect
Emotional Discord

What appears common to the four items is the management of emotions, both positive and negative, and the tensions and problems therein.

The *third factor* (Component 3) is a subset of items in Factor One in the pre-study analysis:

Positive Class Experience
Teacher Support
Life Preparedness
Feelings about School
Extent of Friendship

These items seem more tightly focused around school when compared to Factor One in the pre-study results. What is notable is that Positive Affect, Educational Plans, and Parental Support are now no longer involved.

The *fourth factor* (Component 4) is comprised of the remaining items, and is quite different than any factor identified in the pre-study results:

Educational Plans
Parental Support
Interactional Difficulty

While the interpretation of what is common among these three items is less clear, it is possible that the sense of alienation that runs through the variables that measure Interactional Difficulty signals some strain or tension between students and their parents, and that this affects their plans for their future education and their hopes for life.

Relationships Between Test Anxiety, SOS Scales, and Test Performance

Next we turn to the relationship between pre-post changes in test anxiety, the various SOS scales (including the measures of sociodemographic background and class academic level), and test performance among the students within the experimental and control groups, separately. We begin with the results for the experimental group.

Results for the Experimental Group

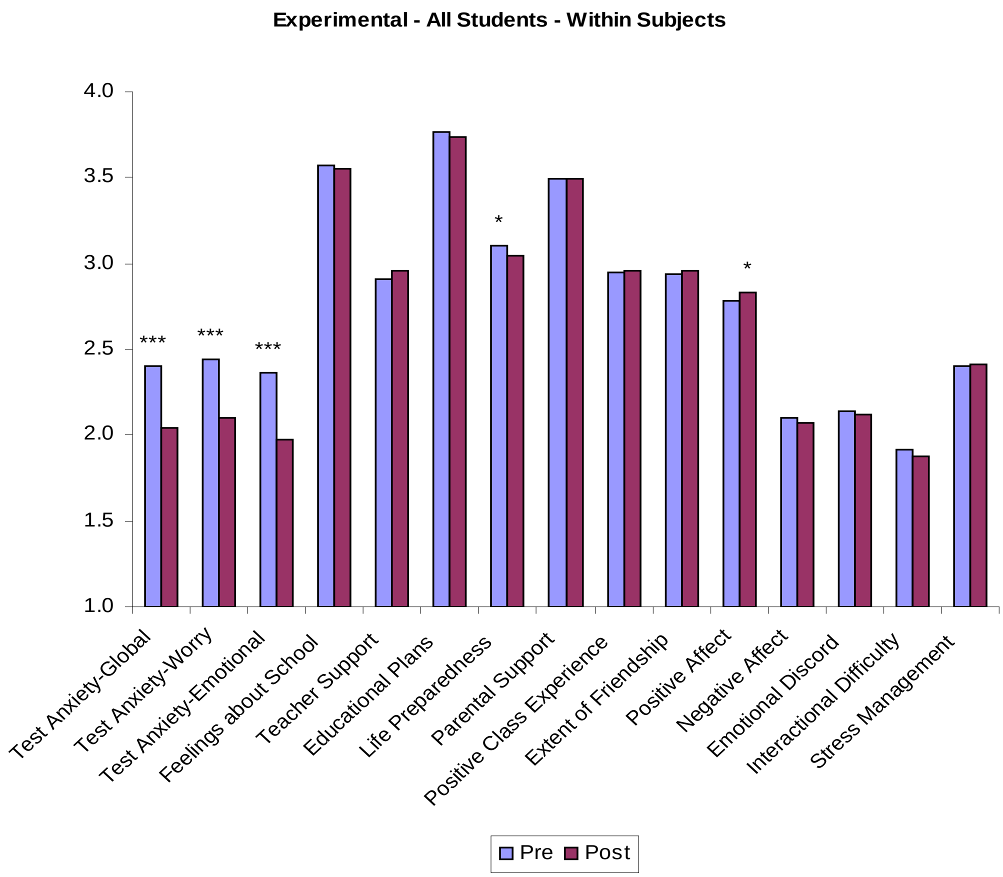
Table VII.3 presents the results of pre-post paired *t*-test comparisons for CST ELA exam performance, test anxiety, and the SOS scales measuring student opinions and emotional dispositions for all students in the experimental group. Although a significant reduction in test anxiety was observed (Test Anxiety-Global: from 2.40 to 2.04, $p < 0.0001$), a small but significant decline in mean CST ELA score (from 348.37 to 345.28, $p < 0.05$) was found. Moreover, a decline was also found on Life Preparedness (from 3.11 to 3.05, $p < 0.05$), and there was a small but significant increase in Positive Affect (from 2.78 to 2.83, $p < 0.05$).

Table VII.3 Pre–Post-Intervention Paired *t*-test Comparisons on CST English-Language-Arts Test Performance, Test Anxiety, and Student Opinion Survey Scales for the Experimental Group

	Pre				Post						
	N	Mean	SD	SEM	N	Mean	SD	SEM	t	df	*p* <
CST English - Language Arts	452	348.37	45.69	2.15	452	345.28	50.80	2.39	-2.20	451	0.05
Test Anxiety-Global	477	2.40	0.90	0.04	477	2.04	0.84	0.04	11.75	476	0.001
Test Anxiety-Worry	476	2.44	0.94	0.04	476	2.10	0.89	0.04	10.25	475	0.001
Test Anxiety-Emotional	472	2.37	0.96	0.04	472	1.97	0.89	0.04	10.72	471	0.001
Feelings about School	488	3.57	0.71	0.03	488	3.56	0.76	0.03	0.52	487	ns
Teacher Support	488	2.91	0.79	0.04	488	2.96	0.81	0.04	-1.31	487	ns
Educational Plans	486	3.77	0.40	0.02	486	3.74	0.46	0.02	1.89	485	ns
Life Preparedness	485	3.11	0.71	0.03	485	3.05	0.74	0.03	2.02	484	0.05
Parental Support	463	3.49	0.67	0.03	463	3.49	0.67	0.03	0.21	462	ns
Positive Class Experience	463	2.95	0.68	0.03	463	2.96	0.74	0.03	-0.18	462	ns
Extent of Friendship	487	2.93	0.71	0.03	487	2.96	0.71	0.03	-0.86	486	ns
Positive Affect	488	2.78	0.63	0.03	488	2.83	0.65	0.03	-2.09	487	0.05
Negative Affect	488	2.10	0.71	0.03	488	2.07	0.70	0.03	1.26	487	ns
Emotional Discord	485	2.14	0.74	0.03	485	2.12	0.76	0.03	0.76	484	ns
Interactional Difficulty	486	1.91	0.62	0.03	486	1.88	0.61	0.03	1.37	485	ns
Stress Management	484	2.41	0.63	0.03	484	2.41	0.66	0.03	-0.14	483	ns

Paired T-test

Figure VII.3 Bar Graph Display of the Results in Table VII.3



We move next to the results of the breakdowns by Gender, Ethnicity, Family Composition, and Class Level (Table VII.4).[10] For Gender, three of the SOS scales emerge with significantly higher scores for Females over Males: Extent of Friendship (3.09 versus 2.84, p <0.0001), and both Negative Affect (2.15 versus 1.99, p <0.001) and Emotional Discord (2.06 versus 2.18, p <0.05). For Ethnicity, the significantly better test performance of Whites and Other Ethnic Groups over Hispanics (352.07 and 347.25, versus 340.27, respectively; p <0.001) is further related to differences by Gender, Family Composition, and Class Level, and to certain SOS scales (Life Preparedness, Extent of Friendship, Feelings About School, Positive Class Experience, and Stress Management), as shown in the Two-Way ANCOVA results presented in Appendix 1 (see Table A.1.1).

[10] The results by ethnicity are provided in Appendix 1 (see Tables A.1.1 and A.1.2), where the results of a more extensive analysis are presented.

For Family Composition, we see that students who live in Other Parent situations had a significantly lower level of test anxiety when compared to those in a Two (biological) Parent family (Test Anxiety-Global: 1.97 versus 2.08, p <0.05). On Class Academic Level, aside from the expected difference in higher CST ELA mean score for students in an Advanced Class over those in a Regular Class (365.88 versus 338.73, p <0.001), only for Feelings about School was there any significant difference, which again was higher for the former (3.67 versus 3.52, p <0.05). The marginally significant result for Teacher Support (not shown: 2.89 versus 3.00, p = 0.06) we disregarded due to the likelihood of a Type 1 error of inference.

Table VII.4 ANCOVA (Significant Results) of Test Anxiety, Test Performance, and Selected SOS Scales by Gender, Family Composition, and Class Level for the Experimental Group

	N	Mean	SD	Adj. Mean	SEM	N	Mean	SD	Adj. Mean	SEM	Mean Sq	F	p <
Gender	Male					Female							
Extent of Friendship	250	2.62	0.71	2.84	0.03	234	3.32	0.50	3.09	0.04	5.43	21.69	0.001
Negative Affect	250	1.92	0.63	1.99	0.03	235	2.22	0.72	2.15	0.03	2.89	10.77	0.001
Emotional Discord	247	1.97	0.70	2.06	0.04	235	2.28	0.78	2.18	0.04	1.78	5.27	0.05
Family Composition	2 Biological parents					Other parents							
Test Anxiety-Global	297	2.08	0.83	2.08	0.04	180	1.97	0.84	1.97	0.04	1.49	4.21	0.05
Test Anxiety-Worry	296	2.14	0.88	2.16	0.04	180	2.04	0.89	2.02	0.05	2.23	5.53	0.05
Class Level	Regular classes					Advanced Classes							
CST English - Language Arts 10	343	327.45	41.15	338.73	1.61	109	401.36	35.07	365.88	3.16	40534.50	51.43	0.001
Feelings about School	375	3.50	0.76	3.52	0.03	113	3.73	0.72	3.67	0.06	1.98	5.83	0.05

ANCOVA

Results for the Control Group

Now we turn to the within-groups results for the control group with respect to the relationship between pre-post changes in test anxiety, test performance, and the various SOS scales, and the measures of sociodemographic background and class academic level.

Table VII.5 presents the results of pre-post paired t-test comparisons. A small but significant reduction on all three measures of test anxiety was observed (e.g., Test Anxiety-Global: from 2.38 to 2.27, p <0.01), and a similar increase in Stress Management was found when pre and post measurements were compared (from 2.32 to 2.42, p <0.01).

Table VII.5 Pre–Post-Intervention Paired *t*-test Comparisons on CST English-Language-Arts Test Performance, Test Anxiety, and Student Opinion Survey Scales for the Control Group

	Pre				Post						
	N	Mean	SD	SEM	N	Mean	SD	SEM	t	df	*p* <
CST English - Language Arts	251	377.47	57.36	3.62	251	380.44	57.55	3.63	-1.55	250	ns
Test Anxiety-Global	255	2.38	0.87	0.05	255	2.27	0.90	0.06	2.73	254	0.01
Test Anxiety-Worry	255	2.30	0.90	0.06	255	2.19	0.90	0.06	2.46	254	0.05
Test Anxiety-Emotional	249	2.45	0.96	0.06	249	2.35	1.01	0.06	2.18	248	0.05
Feelings about School	261	3.74	0.70	0.04	261	3.73	0.77	0.05	0.17	260	ns
Teacher Support	261	2.96	0.78	0.05	261	3.04	0.79	0.05	-1.82	260	ns
Educational Plans	259	3.79	0.37	0.02	259	3.76	0.45	0.03	1.48	258	ns
Life Preparedness	261	2.97	0.68	0.04	261	2.94	0.66	0.04	0.73	260	ns
Parental Support	246	3.55	0.60	0.04	246	3.52	0.64	0.04	1.02	245	ns
Positive Class Experience	245	2.85	0.69	0.04	245	2.79	0.72	0.05	1.42	244	ns
Extent of Friendship	258	3.08	0.63	0.04	258	3.04	0.65	0.04	1.31	257	ns
Positive Affect	261	2.76	0.60	0.04	261	2.74	0.65	0.04	0.40	260	ns
Negative Affect	261	2.19	0.67	0.04	261	2.22	0.73	0.04	-0.93	260	ns
Emotional Discord	260	2.21	0.72	0.04	260	2.27	0.79	0.05	-1.31	259	ns
Interactional Difficulty	261	1.94	0.58	0.04	261	1.97	0.60	0.04	-0.89	260	ns
Stress Management	261	2.32	0.62	0.04	261	2.42	0.65	0.04	-2.85	260	0.01

Paired Sample *t*-test

Moving to the results of the breakdowns by Gender, Ethnicity, Family Composition, and Class Level (Table VII.6),[11] few statistically significant findings are apparent. For Gender, Females report higher test anxiety (Test Anxiety-Global: 2.36 versus 2.14, $p < 0.01$), a greater Extent of Friendship (3.11 versus 2.91, $p < 0.01$), and greater Emotional Discord (2.34 versus 2.15, $p < 0.05$) than Males. On Ethnicity, while Asian students report the greatest Extent of Friendship (3.18 compared to 3.04 and 2.93 for Whites and Other Ethnic Groups, respectively; $p < 0.05$), those in the Other Ethnic Group category had the highest level of Interactional Difficulty (2.10 compared to 1.92 for both the White and Asian categories; $p < 0.05$). And while there were no significant differences for the two categories of Family Composition, students in an Advanced Class scored higher, as expected, on the CST ELA (385.69 versus 367.27, $p < 0.001$) and had higher ratings for Teacher Support (3.13 versus 2.93, $p < 0.01$) and Educational Plans (3.82 versus 3.68, $p < 0.001$) than their classmates in a Regular Class. Consistent with our policy on Type 1 inference errors, we viewed the marginally significant result for Interactional Difficulty (not shown: 1.94 versus 2.04, $p = 0.07$) as due to chance.

[11] The results for ethnicity are provided in Appendix 1 (Tables A.1.3 through A.1.5.

Table VII.6 ANCOVA (Significant Results) of Test Anxiety, Test Performance, and Selected SOS Scales by Gender and Class Level for the Control Group

	N	Mean	SD	Adj. Mean	SEM	N	Mean	SD	Adj. Mean	SEM	Mean Sq	F	*p* <
Gender	Male					Female							
Test Anxiety-Global	99	1.94	0.85	2.14	0.06	150	2.49	0.87	2.36	0.05	2.72	8.15	0.01
Test Anxiety-Worry	99	1.87	0.82	2.01	0.07	150	2.41	0.90	2.32	0.05	5.59	13.88	0.001
Extent of Friendship	100	2.69	0.66	2.91	0.05	153	3.26	0.54	3.11	0.04	1.86	7.94	0.01
Emotional Discord	101	2.02	0.69	2.15	0.06	153	2.43	0.81	2.34	0.05	2.16	5.81	0.05
Class Level	Regular classes					Advanced Classes							
CST English - Language Arts 10	112	338.64	51.20	367.27	3.05	139	408.76	40.54	385.69	2.68	13814.03	17.07	0.001
Teacher Support	119	2.85	0.84	2.93	0.05	142	3.20	0.71	3.13	0.05	2.38	6.83	0.01
Educational Plans	117	3.63	0.57	3.68	0.03	142	3.86	0.27	3.82	0.03	1.35	10.34	0.001

ANCOVA

Summary of Findings

The results of the factor analysis brought to light some interesting commonalities and differences in the latent structure of student perceptions between the two schools:

- Overall, both groups had a somewhat similar factor structure both in terms of number of principal components (3 for the experimental group and 4 for the control group) and categorization of item content
- The primary themes of the factor categories were: emotions and relations; school-related perceptions, feelings, and relations; and test anxiety
- The most notable pre-post difference was the ranking of the factor components (in terms of statistical power): test anxiety went from being ranked as the second component in both groups at baseline, to becoming the least salient component in the experimental group and the most salient component in the control group by the time of post-study measurement
- Another pre-post difference involved the ranking of the factor involving school-related perceptions, feelings, and relations: it went from being the least salient to the most salient factor in the experimental group, and moved from being the most salient to the third lowest in salience in the control group.

The ANCOVA results reveal some additional pre-post similarities and differences between the students in the two groups:

- Common to both groups was a decline in mean test anxiety.
- In both groups, females had higher levels of test anxiety and more emotional discord than males; yet they also had more extensive friendships than males.
- In both groups, students in an advanced class had better test performance on the CST ELA than those in a regular class.
- The main difference was the decline in test performance on the CST ELA in the experimental group; within the experimental group, Hispanics had significantly lower mean test scores than students in the White and Other Ethnic Group categories.

A summary table of significant results from the within-groups analyses of the experimental and control groups is provided in Appendix 2.

Chapter VIII
Comparison of Experimental and Control Groups Using Total Samples

This chapter begins the first of a series which presents results of the data on the outcomes of the TestEdge intervention. We begin our investigation of the intervention effects with an analysis of the equivalence of the experimental and control groups in relation to certain basic characteristics. Following an analysis which shows a relationship between how frequently students practiced the TestEdge tools and their increased use of the tools across different life situations, we present the results of an analysis of pre-post changes in test anxiety which shows a notable difference between the two schools by the end of the study. This is followed by a bivariate analysis of pre-post changes in test performance and the SOS scales in the two groups, and an examination of the intervention effects on test anxiety change and test performance change.

Analysis Considerations

For the various analyses reported in this section, all of the cases in the experimental group and control group were used with the exception of those excluded under the "list-wise deletion" of cases with missing values on variables included in the analysis at hand. In the following section (Chapter IX), a different analysis strategy is pursued in which we use multivariate techniques to seek evidence of the effects of the TestEdge intervention in subpopulations within the experimental group sample. To clarify the distinction between these two different levels of analysis, we use the term *"total samples analysis"* to refer to our use of the whole samples of the experimental group and control group for what follows in this section. We use the term *"sub-samples analysis"* for our investigation of different groupings of cases within the experimental and control groups when matched on various sociodemographic variables and/or SOS scales, the results of which are presented in Chapter IX.

Baseline Equivalence

A key requirement of the quasi-experimental design is that the experimental and control groups are equivalent, or closely matched, on all sociodemographic variables, SOS scales, test anxiety, and test performance at the moment of baseline measurement. Any difference observed between the experimental group and control group on these factors raises the possibility that changes observed in test anxiety or in test performance may be explained, at least in part, by this difference.

Table VIII.1 and Figure VIII.1 present the results of baseline comparisons using a single-factor analysis of variance. Importantly, there is no significant difference between the experimental and control schools on all three measures of test anxiety (TAI-Global, 2.40 versus 2.38, *p* ns; TAI-Worry, 2.43 versus 2.31, *p* ns; and TAI-Emotional, 2.37 versus 2.47, *p* ns, respectively). However, four factors were identified for which there was a significant baseline difference between the two schools. The first is test performance.[12] On the 9th grade CST English-Language Arts test (ELA9), the students in the control group outperformed those in the experimental group by a notable margin—a difference in mean score of 33 points (ELA9 mean score: 380.62 versus 348.34, $p < 0.001$). The means for Feelings about School and for Extent of Friendship are also higher in the control group (3.74 versus 3.57, $p < 0.01$; and 3.08 versus 2.93, $p < 0.01$). But Life Preparedness is significantly higher in the experimental group (3.11 versus 2.97, $p < 0.01$). With the exception of these four factors, there were no other significant differences at baseline between the two schools.

[12]The reader is reminded that due to the various combinations of mathematics subjects and tests taken by different groups of students in the 9th and 10th grades, we do not have a common measure of test performance in mathematics that is standard for all students in the sample.

It is likely that the lack of match between the two schools in socioeconomic level and ethnic composition, noted in Chapter I, accounts, at least in part, for these observed differences. While, given this lack of match, the baseline difference in test performance is not unexpected, it has implications for our analysis strategy: to find evidence of the intervention effects requires matching students from the two schools who have comparable 9th grade test scores to control for this difference. This is addressed in Chapter IX.

In addition to the ANOVA just presented, we also examined the baseline levels of test anxiety by intervention status when broken down by gender, ethnicity, family composition, and class academic level (Table VIII.2).

Using the same three categories of the Test Anxiety-Global scale as was used above for the analysis of the whole sample presented in Table VI.3—viz., Low test anxiety (TAI-Global scale score of <2.0, less than "Sometimes"), Medium test anxiety (>2.1 - <3.0, more than "Sometimes" but less than "Often"), and High test anxiety (>3.0, "Often" or "Almost Always")—the table shows that while the pattern is the same on family composition and class academic level for the experimental group and the control group, there is evidence of a relationship with Hispanic ethnicity and gender. Thus while 37.1% of the Hispanic students in the experimental group and 25.8% in the control group had a low level of test anxiety, 29.1% of experimental group and 41.9% of the control group had a high level of test anxiety. However, because of the large difference in case counts (N = 244, experimental group; N = 32, control group), we were not able to investigate what other sociodemographic factors might explain the notable difference in baseline test anxiety observed in these two groups of Hispanic students. It is likely that the higher socioeconomic level of Hispanic students in the control school may account, in part, for this difference. In relation to gender, a similar pattern to that we observed for the whole sample remains relatively invariant across both the experimental and the control groups. In both groups females are much more likely than males to report a high level of test anxiety (16.5% for males compared to 39.3% for females in the experimental group, and 16.8% compared to 28.1%, respectively, in the control group).

Table VIII.1 Baseline Comparisons

	Experimental (N=488)		Control (N=261)			
	Mean	SD	Mean	SD	F	*p* <
CST Math 9	320.81	50.96	354.80	50.93	71.60	0.0000
CST English - Language Arts 9	348.34	45.46	380.62	57.38	67.78	0.001
Test Anxiety-Global	2.40	0.90	2.38	0.87	0.06	ns
Test Anxiety-Worry	2.43	0.94	2.31	0.90	3.29	ns
Test Anxiety-Emotional	2.37	0.96	2.47	0.97	1.86	ns
Feelings about School	3.57	0.71	3.74	0.70	9.61	0.01
Teacher Support	2.91	0.79	2.96	0.78	0.70	ns
Educational Plans	3.77	0.40	3.79	0.37	0.78	ns
Life Preparedness	3.11	0.71	2.97	0.68	7.22	0.01
Parental Support	3.49	0.67	3.56	0.60	1.55	ns
Positive Class Experience	2.95	0.68	2.85	0.68	3.47	ns
Extent of Friendship	2.93	0.71	3.08	0.63	7.26	0.01
Positive Affect	2.78	0.63	2.76	0.60	0.32	ns
Negative Affect	2.10	0.71	2.19	0.67	2.68	ns
Emotional Discord	2.14	0.74	2.21	0.72	1.54	ns
Interactional Difficulty	1.92	0.62	1.94	0.58	0.34	ns
Stress Management	2.40	0.63	2.32	0.62	3.00	ns

Single Factor ANOVA

Figure VIII.1 Baseline Comparisons

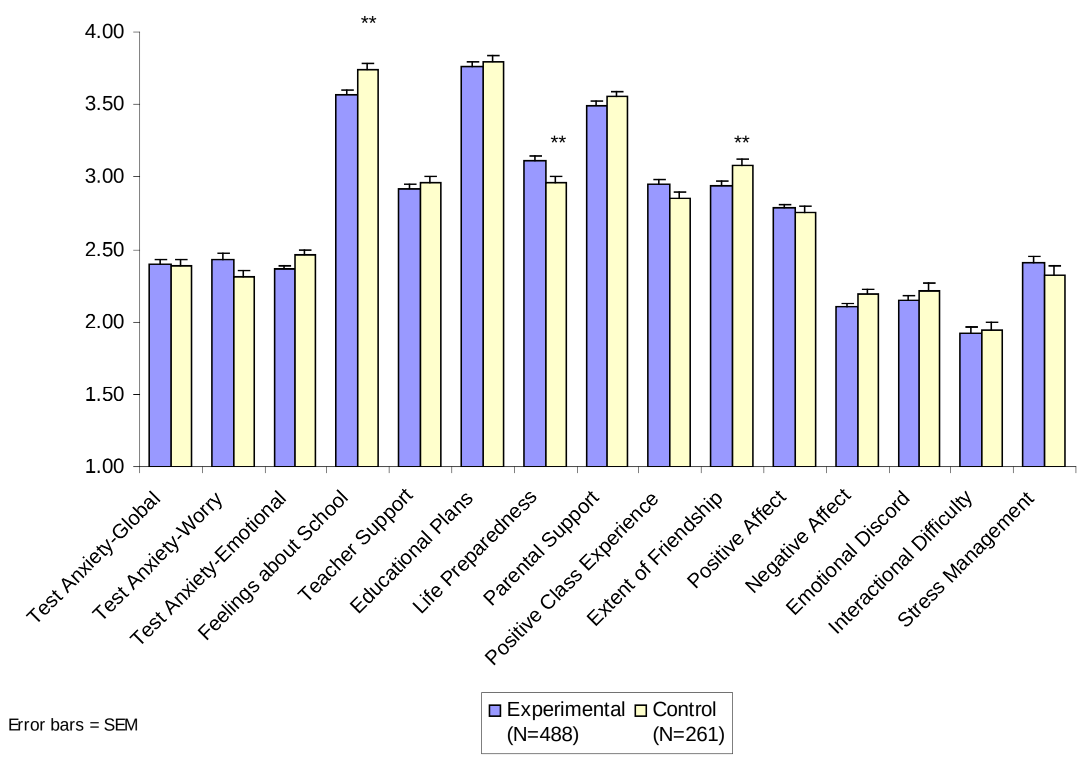


Table VIII.2 Baseline Test Anxiety Levels by Intervention Status and Selected Sociodemographic Characteristics

		Experimental Group Test Anxiety Level				Control Group Test Anxiety Level		
	N	Low	Med	High	N	Low	Med	High
Gender								
Male	248	50.4%	33.1%	16.5%	101	55.4%	27.7%	16.8%
Female	234	26.9%	33.8%	39.3%	153	30.1%	41.8%	28.1%
Ethnicity								
Caucasian	176	40.3%	33.0%	26.7%	138	44.2%	35.5%	20.3%
Hispanic	237	37.1%	33.8%	29.1%	31	25.8%	32.3%	41.9%
Other	64	45.3%	31.3%	23.4%	89	37.1%	39.3%	23.6%
Family Composition								
Both Bio Parents	302	39.4%	32.8%	27.8%	168	40.5%	38.1%	21.4%
Other Parents	181	38.7%	34.3%	27.1%	88	38.6%	34.1%	27.3%
Class Academic Level								
Regular	372	38.4%	32.5%	29.0%	119	37.8%	34.5%	27.7%
Advanced	113	41.6%	36.3%	22.1%	141	41.1%	38.3%	20.6%

Practice and Use of TestEdge Tools in the Intervention School

A key element of the HeartMath program for students in the intervention school was instruction and training by their teacher in use of the emotional management tools in the TestEdge program. After being taught the HeartMath tools, students were encouraged to practice the tools on a daily basis, both in and outside of school. However, the results of interviews with teachers and observations from field workers (see Chapter XIII) suggest that not only did some teachers resent the intrusion of the TestEdge program into class time, but some were also reluctant to spend a few minutes each day on student practice of the TestEdge tools.

To obtain a measure of student practice and use of the Test Edge tools, we surveyed students in the intervention school with the following two questions:

1. *During this school term, have you practiced the TestEdge tools outside class in your own time?*
2. *During this school term, have you used the TestEdge tools to improve your performance or help you in the following situations?*

Data from these questions are presented in Tables VIII.3.a and VIII.3.b.

Table VIII.3.a presents data on the frequency of student practice of the TestEdge tools, both overall and broken down by gender, family composition, ethnicity, and academic class level. While most students (71%) said they used the tools outside class, 29% said that they did not. Just over half (52%) used them either 1–2 times/term or 1–2 times/month, and just under one fifth (19%) used them 1–2 times/week or almost daily. There is little evidence that the frequency of use is related to gender, family composition, ethnicity, or to the academic level of the class in which students belonged.

Moving to the data on the student applications of the TestEdge tools (Table VIII.3b), just over half of the students (52%) indicated that they used the tools at school and on tests; 26% used them in an extracurricular activity (sports, music, etc.); 29% used them in their relationships at school, or with friends and family; 28% used them to solve difficult or complicated problems; and 36% said they used the tools to manage their emotions. There is some evidence that more females than males use the tools at school and on tests, in their relationships, to solve difficult problems, and to manage their emotions. In addition, there is evidence that more students in an advanced level class use the tools to manage their emotions than those in regular classes. Finally, there is somewhat greater use of the tools in various life situations by students in the Other ethnic group.

Table VIII.3.a Frequency of Practice of TestEdge Tools Outside Class —Intervention School

Q16. During this school term, have you practiced using the TestEdge tools outside class in your own time?

	Total	Male	Female	Hispanic	White	Other	2 Bio Parents	Other Partents	Regular	Advanced
No	29%	34%	24%	24%	33%	28%	27%	32%	30%	26%
Yes: 1-2 times this term	37%	29%	45%	42%	34%	26%	40%	31%	36%	37%
Yes: 1-2 times a month	15%	16%	14%	13%	17%	21%	13%	18%	13%	20%
Yes: 1-2 times a week	13%	14%	12%	14%	12%	18%	14%	12%	14%	10%
Yes: Almost every day	6%	7%	5%	7%	4%	7%	6%	7%	6%	7%
Total Count	430	221	206	206	159	57	271	159	327	103

Table VIII.3.b Use of TestEdge Tools in Different Situations —Intervention School

Q17 During this school term, have you used the TestEdge tools to improve your performance or help you in the following situations? Mark all answers that apply.

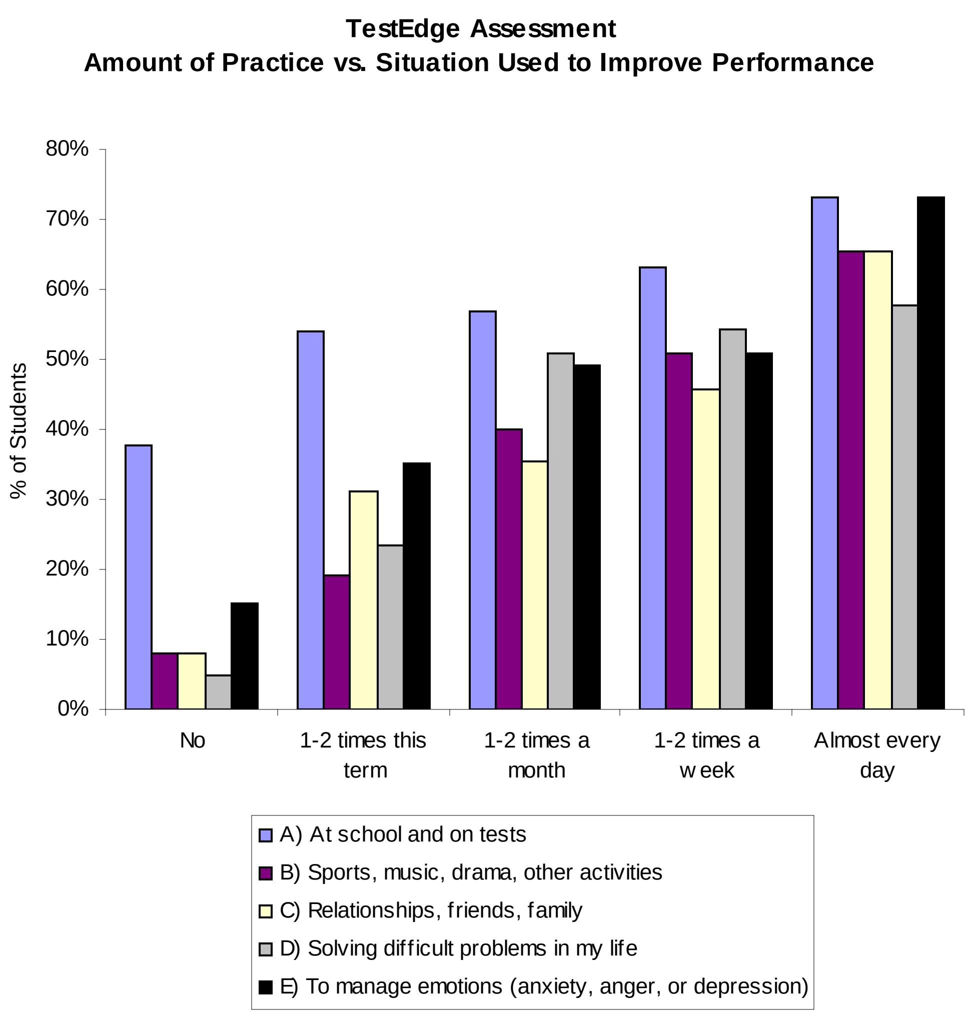
	Total (N=433)		Males (N=224)		Females (N=206)		Regular Classes (N=329)		Advanced Classes (N=104)		2 Bio Parents (N=273)		Other Parents (N=160)		White (N=160)		Hispanic (N=208)		Other (N=57)	
A) At school and on tests	224	52%	103	46%	119	58%	167	51%	57	55%	139	51%	85	53%	85	53%	108	52%	30	53%
B) In sports, music, drama, or other activities	112	26%	64	29%	48	23%	78	24%	34	33%	75	27%	37	23%	46	29%	42	20%	23	40%
C) In my relationships at school, with my friends, or with my family	125	29%	49	22%	76	37%	94	29%	31	30%	77	28%	48	30%	38	24%	62	30%	24	42%
D) To solve difficult or complicated problems in my life	122	28%	56	25%	66	32%	93	28%	29	28%	71	26%	51	32%	44	28%	63	30%	15	26%
E) To manage my emotions--e.g., when anxious, angry or depressed	154	36%	63	28%	90	44%	107	33%	47	45%	100	37%	54	34%	53	33%	73	35%	26	46%

Figure VIII.4 shows that there is a strong positive relationship between the frequency of student use of the tools outside class and the number of different situations in which they use the tools; these situations include school and tests; extracurricular activities such as sports, music, etc.; various interpersonal relationships; solving complicated problems in life; and managing emotions.

However, the results of an ANCOVA (not shown) for all students in the intervention school did not find evidence of a relationship between the frequency of use of the TestEdge tools and pre-post intervention test anxiety reduction. There was also no evidence of a relationship between student use of the tools and 10th grade CST ELA test performance. While the lack of a relationship between tool use and test anxiety is contrary to the direct effects expected from the intervention, until we have the opportunity to investigate this question in depth, we regard this result as preliminary rather than definitive. Among the issues to be investigated in further analysis is the accuracy in the reporting of tool use among a subset of students in the intervention school. Also reserved for future analysis is the degree to which there are sub-samples of students within the intervention group, distinguished by sociodemographic or other characteristics, for which the expected relationship between tool use and test anxiety is observed.

Returning to other results from the ANCOVA (not shown), there was evidence of a relationship between frequency of tool use and pre-post intervention change on three of the SOS scales. On Life Preparedness—the degree to which students believe what they have learned at school prepares them for life—there is a significant difference associated with frequency of tool use (from 2.87 for "none," 3.05 for 1–2 times this term or 1–2 times a month, to 3.27 for 1–2 times a week or almost every day; p <0.05 for all three pair-wise combinations). There is also a significant increase in reports of a Positive Class Experience comparing students who report no tool use with those who say they use the tools ≥ 1–2 per week (2.88 versus 3.12, p <0.05). And finally, there is a significant increase in reports of better Stress Management skills when the two categories of tool use are compared to those students who report no tool use (2.31 versus 2.39, and 2.31 versus 2.61; p <0.001 for both paired comparisons). In short, there is a positive relationship between use of the TestEdge tools, feeling better prepared for life, having a more positive class experience, and having a greater ability to manage stress.

Figure VIII.4 Relationship Between Frequency of Practicing TestEdge Tools and Tool Use in Different Situations



Changes in Test Anxiety Over Time

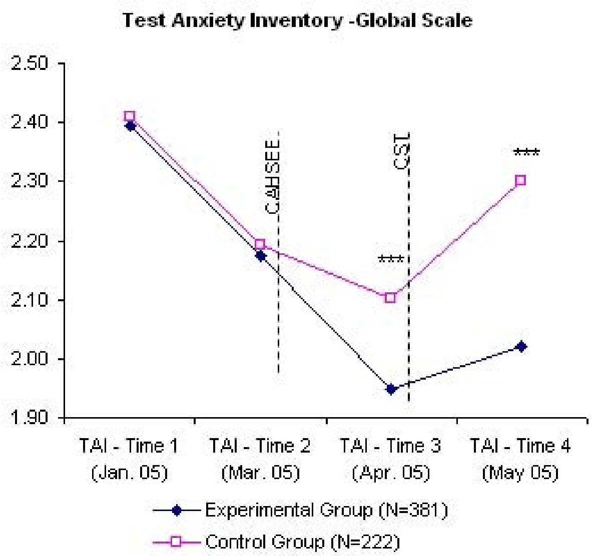
Student test anxiety was measured at four different times in the study—the first measurement was obtained in the pre-study survey administration, in January; the second in March, two weeks before students took the CAHSEE; the third two weeks before students took the CST, in April; and the fourth at the end of the study in the post-study survey administration, in May. A breakdown of the changes in test anxiety by intervention status is presented in Table VIII.5 and Figure VIII.5.

Viewing the over-time trajectories graphically displayed in Figure VIII.5 (next page), we can see that the experimental and control groups both begin with virtually the same level of test anxiety (2.40 and 2.41, respectively; *p* ns) at Time 1, which declines to virtually the same level (2.17 and 2.19, *p* ns) at Time 2, after which it further declines to its lowest respective levels and also to separate by a significant margin at Time 3 (1.95 versus 2.10, $p < 0.001$), and then rises differentially—by a small amount in the experimental group (2.02, *p* ns), and a significantly larger amount (2.30, $p < 0.01$) in the control group—at the end of the study at Time 4. An important point to note is the experimental group's steeper decline in test anxiety, which rose only slightly in the month or so between the Time 3 and Time 4 measurements. This appears to be compelling evidence of the effect and efficacy of the TestEdge intervention.

Table and Figure VIII.5 Test Anxiety Inventory-Global Scale: Comparison of Experimental and Control Schools on Repeated Measures at Four Time Points

Repeated Measures ANOVA
Descriptive Statistics

	Experimental Group (N=381)	SD	Control Group (N=222)	SD
TAI - Time 1 (Jan. 05)	2.40	0.91	2.41	0.87
TAI - Time 2 (Mar. 05)	2.17	0.77	2.19	0.75
TAI - Time 3 (Apr. 05)	1.95	0.70	2.10	0.76
TAI - Time 4 (May 05)	2.02	0.83	2.30	0.89



Tests of Within-Subjects Effects

Source	df	Mean Sq.	F	*p* <
TIME	3	13.57	82.22	0.001
TIME * SITE_SCH	3	2.25	13.62	0.001
Error(TIME)	1803	0.17		

Tests of Within-Subjects Contrasts
Measure: MEASURE_1

Source	TIME	df	Mean Sq.	F	*p* <
TIME	Level 1 vs. Level 2	1	26.59	81.63	0.001
	Level 2 vs. Level 3	1	14.00	61.47	0.001
	Level 3 vs. Level 4	1	10.13	36.70	0.001
TIME * SITE_SCH	Level 1 vs. Level 2	1	0.01	0.02	ns
	Level 2 vs. Level 3	1	2.50	11.00	0.001
	Level 3 vs. Level 4	1	2.25	8.14	0.01
Error(TIME)	Level 1 vs. Level 2	601	0.33		
	Level 2 vs. Level 3	601	0.23		
	Level 3 vs. Level 4	601	0.28		

Pre- and Post-Study Bivariate Comparisons

In this section we report the results of pre and post analysis of variance of the thirteen SOS scales and test performance, using intervention status as the independent or grouping variable. To aid the reader, the data are presented in both tabular and graphic form.

Using an analysis of covariance (ANCOVA) to adjust for differences in means at baseline, Table VIII.6 and Figure VIII.6 show the relationship between the intervention and the measures of test anxiety, test performance, and the SOS scales by the end of the study. The first point to note is that the mean pre-post change difference in CST ELA test scores between the two groups is small, at 3.87 points (355.47 for the experimental group versus 359.34 for the control group, p ns). Given the baseline 9th grade CST ELA results reported above, it would appear that the students in the intervention school have closed the test performance gap on their classmates in the control school. However, the narrowing in 10th grade CST ELA test scores is due to the ANCOVA adjustment for baseline differences between the two groups of students. It is likely that the lack of an observed difference in test performance match in the experimental group over the control not only reflects the differences between the two schools in socioeconomic level and ethnic composition, as previously noted, but may also result from a less-than-optimal implementation of the TestEdge program in the intervention school. However, as we will see in the next chapter, when these differences are statistically controlled by using a matched-groups analysis approach, there are a number of subsamples of students within the experimental group for which the expected effects on test anxiety and test performance were observed.

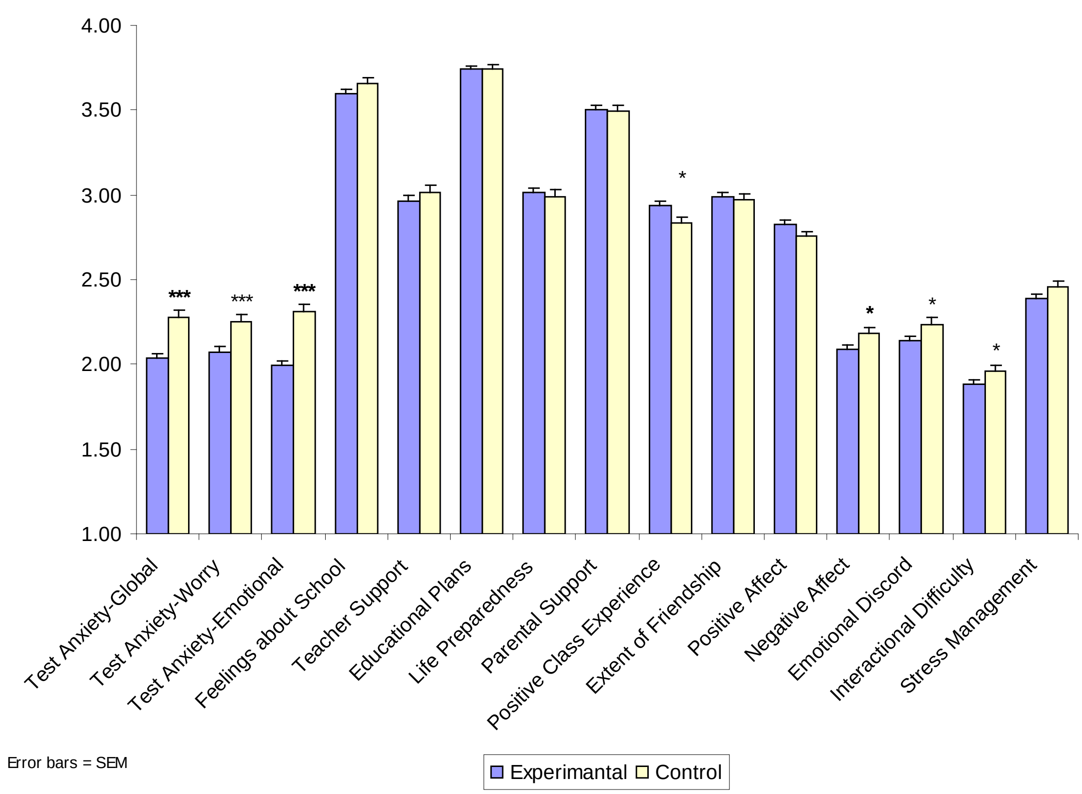
Returning to the ANCOVA results, an important result is that a significantly lower level of test anxiety, after adjusting for baseline differences, was observed in the experimental than in the control group by the end of the study—on both the global measure of test anxiety (TAI-Global, 2.03 versus 2.28, $p < 0.001$) and on each of the Worry (2.07 versus 2.25, $p < 0.001$) and Emotionality (1.99 versus 2.31, $p < 0.001$) subscales. A third result of interest is that the experimental group has a higher mean on Positive Class Experience than the control group (2.94 versus 2.83, $p < 0.05$), and also has a lower mean on Negative Affect, Emotional Discord, and Interactional Difficulty than the latter (2.09 versus 2.19, 2.14 versus 2.24, and 1.89 versus 1.96, respectively; all $p < 0.05$). Both the notable reduction in test anxiety and the modest changes in feelings and emotions just mentioned are consistent with the expected effects of the TestEdge intervention in the experimental group.

Table VIII.6 Post-Study Results Comparing the Experimental and Control Schools Using ANCOVA Adjusted Means

	Experimental			Control					
	Est Mean	SEM	N	Est Mean	SEM	N	Mean Sq.	F	*p* <
CST English - Language Arts 10	355.47	1.41	452	359.34	1.92	251	2215.65	2.54	ns
Test Anxiety-Global	2.03	0.03	477	2.28	0.04	255	10.23	28.47	0.001
Test Anxiety-Worry	2.07	0.03	476	2.25	0.04	255	5.32	12.84	0.001
Test Anxiety-Emotional	1.99	0.03	472	2.31	0.04	249	16.89	34.78	0.001
Feelings about School	3.60	0.03	488	3.66	0.04	261	0.58	1.70	ns
Teacher Support	2.97	0.03	488	3.02	0.04	261	0.42	1.05	ns
Educational Plans	3.74	0.02	486	3.75	0.02	259	0.00	0.01	ns
Life Preparedness	3.02	0.03	485	2.99	0.04	261	0.11	0.33	ns
Parental Support	3.50	0.02	463	3.50	0.03	246	0.01	0.06	ns
Positive Class Experience	2.94	0.03	463	2.83	0.04	245	1.69	4.63	0.05
Extent of Friendship	2.99	0.02	487	2.97	0.03	258	0.07	0.26	ns
Positive Affect	2.83	0.02	488	2.76	0.03	261	0.90	3.57	ns
Negative Affect	2.09	0.02	488	2.19	0.03	261	1.62	5.59	0.05
Emotional Discord	2.14	0.03	485	2.24	0.04	260	1.64	4.59	0.05
Interactional Difficulty	1.89	0.02	486	1.96	0.03	261	0.96	4.10	0.05
Stress Management	2.39	0.02	484	2.46	0.03	261	0.71	2.49	ns

ANCOVA

Figure VIII.6 Post-Study ANCOVA Adjusted Means



Explaining Change in Test Anxiety

We now turn the question of accounting for the change in test anxiety. Two measures of change were examined: that observed between the pre-study baseline measurement and the post-study measurement at the end of the study, and that observed between the pre-study baseline measurement and Time 3, two weeks before students took the CST.

For the pre-post study changes, it will be recalled from the results just presented that in the experimental group the mean reduction in global test anxiety was significantly greater than in the control group. Two different multivariate analyses were conducted in an effort to explain this difference. The first involved a multiple regression analysis, conducted separately on each group, to identify whether the pre-post change on any within-groups factors was significantly related to the decline in test anxiety. The second involved the use of discriminant function analysis in a between-groups approach to investigate how powerful, in relation to the pre-post changes observed for all other factors, the change in test anxiety was as a common trait that statistically distinguished between the two groups.

Starting with the regression analysis, the results (not shown) for both groups were not successful. In the experimental group, two factors—Emotional Discord and Gender—were entered into the stepwise regression model, and together they explain only about 6% of the change in test anxiety (adjusted $R^2 = 0.057$, standardized beta coefficients = 0.207 and -0.120, respectively). In the control group only one factor—Emotional Discord—met the stepwise criteria for inclusion in the statistical model, and it explained only about 5% of the change in test anxiety (adjusted $R^2 = 0.054$, standardized beta coefficient = 0.247).

For the discriminant function analysis, intervention status (experimental versus control group) was used as the nominal dependent variable and the sociodemographic variables (following transformation into binary dummy variables) and SOS scales were treated as independent variables. The results (not shown) showed that two factors—pre-post change in test anxiety and Gender—met the stepwise criteria for inclusion in the multivariate additive model. However, the reduction in Wilks' Lambda (the measure of association between the two groups) was negligible (from 0.959 to 0.947). The resulting canonical discriminant function had both limited statistical power (Eigenvalue = 0.057) and limited predictive utility (59.7% correct classification for predicted group membership; prior probability = 54.1%).

Explaining Change in Test Performance

The small degree of between-groups change observed above in student performance on the CST in ELA (3.87 points) means that there is insufficient change to explain statistically. A similar situation holds for the within-groups changes in student scores on this test from the 9th to the 10th grades. The pre-post difference in mean test scores for each group on the CST ELA is only -3.09 points for the experimental group and -2.97 points for the control group. Again, there is little within-groups change in test performance to explain.

However, as we will see in the next section, a different picture emerges when multivariate techniques are used to construct matched-group comparisons of subsamples within the experimental and control groups in which a certain combination of variables is held constant while intervention status varies. The matched-group comparisons identify a number of subgroups within the experimental school for which significant differences on test anxiety and test performance are found.

Summary of Findings

In seeking evidence of the intervention effects by comparing the full samples of students in the experimental and control groups, the following findings emerged:

- While generally there was evidence of baseline equivalence between the two schools on most factors, there was higher academic performance and a more positive school climate in the control group.
- For students in the experimental group, there was evidence of a strong positive relationship between frequency of TestEdge tool use and the increased use of tools across different life situations; however, there is no evidence of a relationship between reported frequency of tool use and test anxiety reduction.
- There is evidence of a marked pre-post decline on all three measures of test anxiety in the experimental group when compared to the control group.
- There is also corroborating evidence of pre-post declines on measures of negative emotions and relational problems, and a corresponding increase in positive emotions favoring the experimental group over the control group.

- When baseline differences in 9th grade CST ELA test scores are controlled, there is little evidence of a pre-post difference in test performance between the two groups.
- The effort to explain the pre-post changes in test anxiety, using multiple regression and discriminant function analysis within each group and between the two groups, respectively, was not successful.

Chapter IX
Identifying Intervention Effects by Comparing Matched-Group Sub-Samples

In this section we change the level of analysis from that of the total sample of students in each of the experimental and control groups to one of an investigation of intervention effects by comparing various sub-samples within the experimental and control groups. We use multivariate techniques to construct matched-group comparisons on various combinations of sociodemographic variables and other matched-group comparisons on Mathematics subject matter content and English-Language-Arts test scores. This enables the employment of statistical controls in which a certain combination of variables is held constant in a given matched-group sub-sample pair, while the intervention effects on test anxiety and test performance are determined by comparing the experimental group sub-sample with the corresponding control group sub-sample.

In relation to the question of intervention effects, the results from the analysis of total samples in the previous chapter (Chapter VIII) offer, at best, only partial evidence. This is because while there was evidence of a significantly greater decline of test anxiety in the experimental group when compared to the control group, there was no evidence of a difference between the two groups on test performance. For evidence of a fully successful TestEdge intervention, the evidence must show that the pre-post change in test anxiety is directly associated with the expected pre-post change in test performance.

This evidence emerges when we analyze sub-samples from the two schools matched on certain sociodemographic factors and other sub-samples matched on 9th and 10th grade Mathematics subject matter content. This matched-groups approach not only reveals a number of sub-samples in the experimental group for which a reduction in test anxiety was associated with improved emotional disposition, but it also reveals significantly higher test performance in four sub-samples in the experimental group. As will become apparent, the four sub-samples are:

1. White females in a regular academic level class (Regular White Females).
2. Other ethnicity females in an advanced academic level class (Advanced Other Females).
3. Math Group 1 (students who took 9th grade Geometry and 10th grade Algebra 2).
4. Math Group 2 (students who took 9th grade Algebra 1 and 10th grade Geometry).

In two of these sub-samples in the experimental group—Regular White Females and Math Group 1—we found that a post intervention improvement in test performance was associated both with a reduction in test anxiety and with improvement in emotional disposition, as measured by the SOS scales. However, there is also evidence of significantly higher test performance in two sub-populations in the control group, primarily involving males.

Degree of Overlap Among Matched Sub-Samples

Table IX.1 shows that while Math Group 1 and Math Group 2, and Regular White Females and Advanced Other Females sub-populations are each mutually exclusive groups of students within the experimental group, there is some degree of overlap with the two other sub-populations. The degree of overlap in descending order was 11% for Regular White Females (i.e., 89% belonged only to this sub-population), 16% for Math Group 2 (84% belonged exclusively to this sub-population), 22% for Math Group 1 (78% belonged exclusively to this sub-population), and 62% for Advanced Other Females (38 % belonged exclusively to this sub-population). While there has not yet been time to investigate the influence of the degree of overlap among these groups in the results that follow, this will be an important objective in future analysis.

Table IX.1 Degree of Overlap Among Matched Sub-Samples from the Experimental Group

	Math group 1*	Math group 2*	Regular White Females	Advanced Other Females
Math group 1	100	0	5	24
Math group 2	0	56	4	7
Regular White Females	5	4	73	0
Advanced Other Females	24	7	0	19
Total N	129	67	82	50

	Math group 1*	Math group 2*	Regular White Females	Advanced Other Females
Math group 1	78%	0%	6%	48%
Math group 2	0%	84%	5%	14%
Regular White Females	4%	6%	89%	0%
Advanced Other Females	19%	10%	0%	38%
Total N	129	67	82	50

	9th grade	10th grade
Math group 1	Geometry	Algebra 2
Math group 2	Algebra 1	Geometry

Sub-Samples Matched on Sociodemographic Factors and Class Academic Level

Before proceeding, a distinction must be made between Class Academic Level, one of the variables used in the analyses that follow, and Class Mean Test Performance, a key variable used in a subsequent section. Class Academic Level is the nominal classification made by the two high schools to group students into classes of different academic levels on the basis of a student's academic performance. However, Class Mean Test Performance is based on the actual mean performance of students in a given class on the 9th grade CST, as will become clear below.

Change in Test Anxiety

We began this investigation of sub-samples by examining the degree to which the observed changes in test anxiety from baseline through Time 3 were related to sociodemographic factors. For this investigation an ANCOVA was performed in which the test anxiety delta at Time 3 was used as the dependent variable while the categories in the Gender, Ethnicity, Family Composition, and Class Academic Level variables were used as independent variables; intervention status (experimental or control group) was used as the grouping variable. The results are shown in Table IX.2 and Figure IX.2.

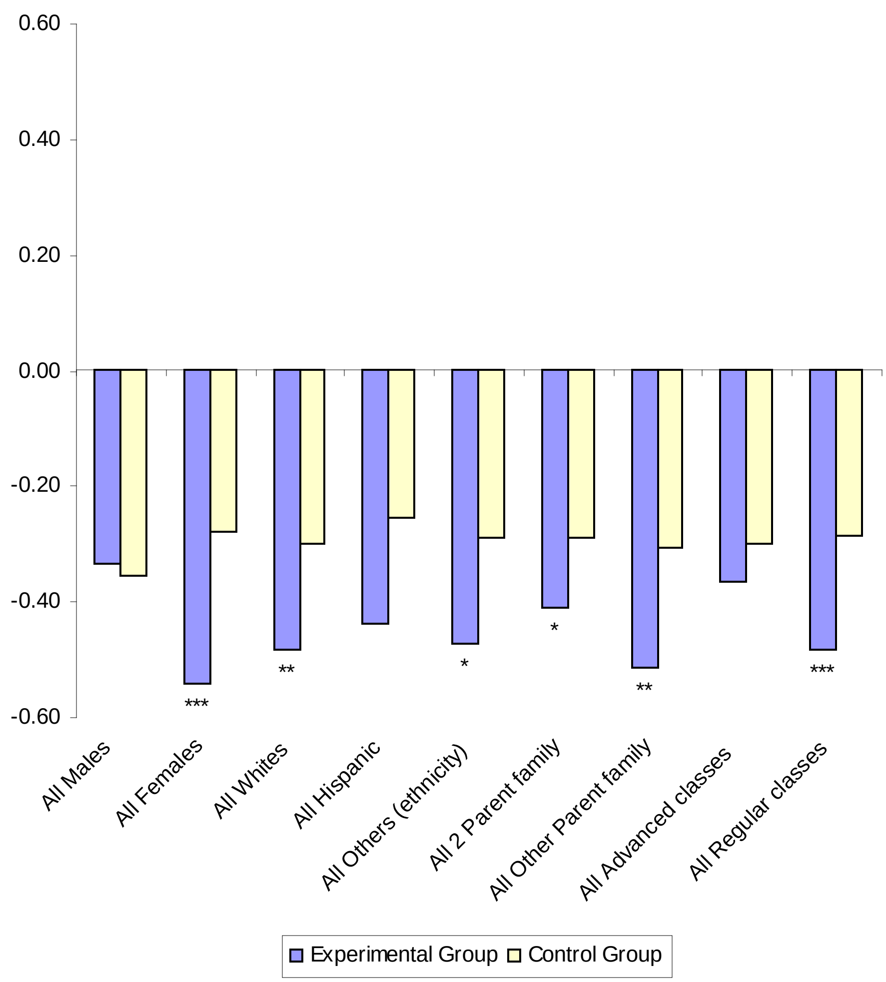
The most striking result is that, with only three exceptions (All Males, All Hispanic, All Advanced Class), six characteristics are each associated with a significantly greater reduction in test anxiety by Time 3 in the experimental group than in the control group. The six are: All Females (-0.54 versus -0.28, $p < 0.001$), All Whites (-0.48 versus -0.30, $p < 0.01$), All Other Ethnic Groups (-0.47 versus -0.29, $p < 0.05$), All those in a Two biological Parent family (-0.41 versus -0.29, $p < 0.05$), All those in some Other family (-0.51 versus -0.31, $p < 0.01$), and All Regular Classes (-0.48 versus -0.29, $p < 0.0001$). The last result is noteworthy because most students are *not* placed in advanced academic level classes, and it appears that the TestEdge intervention was successful in helping these students achieve reductions in test anxiety just before they took the CST.

Table IX.2 ANCOVA of Post-Intervention Test Anxiety Change at Time 3 by Gender, Ethnicity, Family Composition, and Class Academic Level by Intervention Status

	Experimental Group					Control Group							
	N	Mean	SD	Adjusted Means	SEM	N	Mean	SD	Adjusted Means	SEM	Mean Sq.	F	*p* <
All Males	205	-0.33	0.63	-0.33	0.03	90	-0.36	0.62	-0.35	0.05	0.03	0.13	ns
All Females	207	-0.56	0.70	-0.54	0.04	145	-0.25	0.56	-0.28	0.04	5.80	20.75	0.001
All Whites	151	-0.49	0.70	-0.48	0.04	125	-0.29	0.63	-0.30	0.05	2.35	8.04	0.01
All Hispanic	202	-0.42	0.66	-0.44	0.04	30	-0.37	0.61	-0.25	0.09	0.85	3.42	ns
All Others (ethnicity)	58	-0.47	0.70	-0.47	0.07	84	-0.29	0.53	-0.29	0.06	1.13	4.35	0.05
All 2 Parent family	258	-0.42	0.68	-0.41	0.03	158	-0.28	0.51	-0.29	0.04	1.45	5.57	0.05
All Other Parent family	155	-0.50	0.68	-0.51	0.04	79	-0.33	0.73	-0.31	0.06	2.22	7.69	0.01
All Advanced classes	97	-0.37	0.67	-0.36	0.05	129	-0.29	0.53	-0.30	0.05	0.24	0.93	ns
All Regular classes	318	-0.47	0.68	-0.48	0.03	112	-0.32	0.65	-0.29	0.05	3.17	11.61	0.0001

ANCOVA

Figure IX.2 Bar Graph of Adjusted Means from Table IX.2



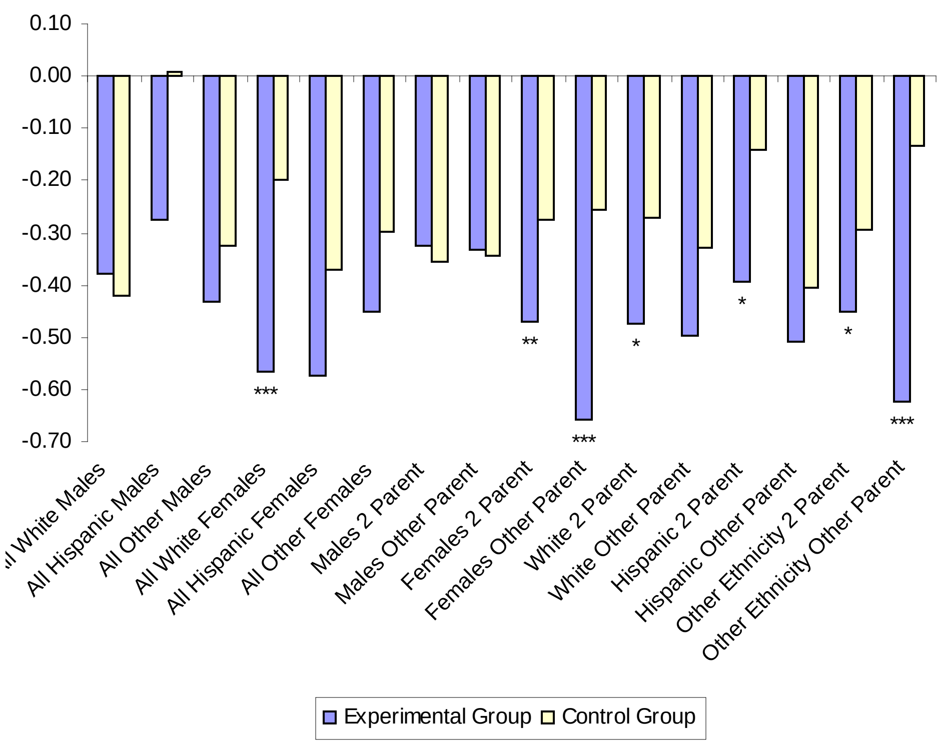
In order to refine the identification of the sub-samples even further, we next used an ANCOVA to investigate the degree to which the observed changes in test anxiety from baseline through Time 3 were related to certain combinations of sociodemographic characteristics and class academic level. Two analyses were conducted using the change in test anxiety by Time 3 as the dependent variable; intervention status was used as the grouping variable. In the first analysis, the different logical bivariate combinations of categories for Gender, Ethnicity, and Family Composition were used as independent variables. In the second analysis, we examined the bivariate combinations of Class Academic Level with Gender, Ethnicity, and Family Composition, separately. The results are shown in Table IX.3 and Figure IX.3, and in Table IX.4 and Figure IX.4.

A number of significant results emerged for certain combinations of characteristics, all of which show greater reductions in test anxiety in the experimental group compared to the control group (Table IX.3 and Figure IX.3). In line with our previous finding of a greater reduction in test anxiety among females, three of these sub-sample combinations involve females—All White Females (-0.57 versus -0.20, $p < 0.001$), Females with Two biological Parents (-0.47 versus -0.28, $p < 0.01$), and Females in Other family situations (-0.66 versus -0.26, $p < 0.001$). Intact family status is the second common factor in three other significant combinations: White with Two biological Parents (-0.48 versus -0.27, $p < 0.05$), Hispanic with Two biological Parents (-0.39 versus -0.14, $p < 0.05$), Other Ethnic Group with Two biological Parents (-0.45 versus -0.30, $p < 0.05$). The last combination showing a significant reduction in test anxiety was the sub-sample of Other Ethnic Groups in Other family situations (-0.62 versus -0.13, $p < 0.001$).

Table IX.3 ANCOVA of Post-Intervention Test Anxiety Change at Time 3 for Matched Sub-Samples on Bivariate Combinations of Gender, Ethnicity, and Family Composition by Intervention Status

	Experimental Group					Control Group							
	N	Mean	SD	Adjusted Means	SEM	N	Mean	SD	Adjusted Means	SEM	Mean Sq.	F	*p* <
All White Males	73	-0.40	0.70	-0.38	0.06	54	-0.39	0.64	-0.42	0.07	0.06	0.22	ns
All Hispanic Males	101	-0.27	0.60	-0.28	0.05	8	-0.08	0.69	0.01	0.17	0.60	2.62	ns
All Other Males	30	-0.40	0.56	-0.43	0.08	28	-0.36	0.57	-0.33	0.08	0.16	0.85	ns
All White Females	77	-0.57	0.70	-0.57	0.06	70	-0.19	0.60	-0.20	0.06	4.92	17.30	0.001
All Hispanic Females	101	-0.57	0.68	-0.58	0.05	20	-0.41	0.52	-0.37	0.12	0.70	2.65	ns
All Other Females	27	-0.52	0.84	-0.45	0.11	55	-0.27	0.51	-0.30	0.07	0.42	1.39	ns
Males 2 Parent	129	-0.33	0.63	-0.33	0.04	57	-0.34	0.57	-0.36	0.06	0.03	0.15	ns
Males Other Parent	75	-0.33	0.65	-0.33	0.06	32	-0.36	0.70	-0.34	0.10	0.00	0.01	ns
Females 2 Parent	128	-0.50	0.72	-0.47	0.05	99	-0.23	0.45	-0.28	0.05	2.08	7.53	0.01
Females Other Parent	79	-0.65	0.67	-0.66	0.06	44	-0.26	0.75	-0.26	0.08	4.58	16.92	0.001
White 2 Parent	92	-0.48	0.67	-0.48	0.05	78	-0.26	0.52	-0.27	0.06	1.73	6.44	0.05
White Other Parent	59	-0.50	0.74	-0.50	0.07	46	-0.33	0.79	-0.33	0.08	0.72	2.23	ns
Hispanic 2 Parent	131	-0.37	0.65	-0.39	0.04	18	-0.28	0.55	-0.14	0.11	0.98	4.23	0.05
Hispanic Other Parent	70	-0.50	0.66	-0.51	0.07	11	-0.44	0.68	-0.41	0.16	0.09	0.33	ns
Other Ethnicity 2 Parent	125	-0.47	0.70	-0.45	0.05	140	-0.28	0.51	-0.30	0.04	1.62	5.95	0.05
Other Ethnicity Other Parent	25	-0.51	0.57	-0.62	0.09	21	-0.27	0.64	-0.13	0.10	2.50	12.63	0.001

Figure IX.3 Bar Graph of Adjusted Means from Table IX.3



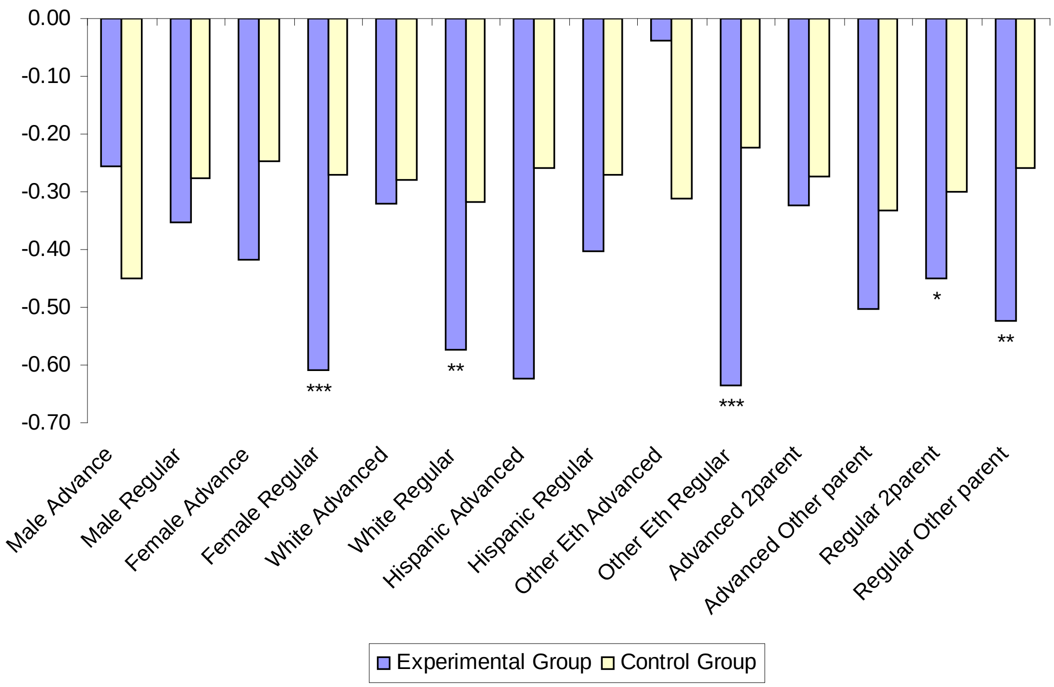
Moving to the ANCOVA of bivariate combinations of Class Level with Gender, Ethnicity, and Family Composition (Table IX.4 and Figure IX.4), five significant results emerge, involving a greater reduction of test anxiety in the experimental group in relation to the control group. Interestingly, *all* of these involve students in a Regular (non-advanced level) Class, and encompass five (71%) of the seven combinations investigated involving this characteristic: Female students in a Regular Class (-0.61 versus -0.27, $p < 0.001$), White students in a Regular Class (-0.57 versus -0.32, $p < 0.01$), students from an Other Ethnic Group in a Regular Class (-0.63 versus -0.22, $p < 0.001$), students in a Regular Class from intact families (-0.45 versus -0.30, $p < 0.05$), and students in a Regular Class from Other family circumstances (-0.53 versus -0.26, $p < 0.01$).

From the perspective of intervention efficacy, this finding is noteworthy because it suggests that the TestEdge program is effective in helping most students in regular classes to manage their test anxiety. It also suggests the program is readily accessible to students of normal academic ability and does not require superior intellectual or learning ability.

Table IX.4 ANCOVA of Post-Intervention Test Anxiety Change at Time 3 for Matched Sub-Samples on Bivariate Combinations of Class Academic Level with Gender, Ethnicity, and Family Composition by Intervention Status

	Experimental Group					Control Group							
	N	Mean	SD	Adjusted Means	SEM	N	Mean	SD	Adjusted Means	SEM	Mean Sq.	F	*p* <
Male Advanced	40	-0.29	0.55	-0.26	0.07	38	-0.41	0.62	-0.45	0.07	0.73	3.54	ns
Male Regular	165	-0.34	0.65	-0.35	0.04	52	-0.31	0.63	-0.28	0.07	0.24	0.95	ns
Female Advanced	57	-0.42	0.74	-0.42	0.07	89	-0.24	0.49	-0.25	0.06	1.01	3.75	ns
Female Regular	150	-0.61	0.68	-0.61	0.04	56	-0.26	0.65	-0.27	0.07	4.62	16.19	0.001
White Advanced	53	-0.30	0.56	-0.32	0.07	65	-0.29	0.55	-0.28	0.06	0.05	0.20	ns
White Regular	98	-0.59	0.75	-0.57	0.06	60	-0.29	0.72	-0.32	0.07	2.41	7.77	0.01
Hispanic Advanced	28	-0.67	0.62	-0.62	0.08	5	0.03	0.38	-0.26	0.18	0.52	3.34	ns
Hispanic Regular	174	-0.38	0.65	-0.40	0.04	25	-0.45	0.62	-0.27	0.10	0.38	1.45	ns
Other Eth Advanced	15	-0.03	0.91	-0.04	0.14	58	-0.32	0.53	-0.31	0.07	0.90	3.03	ns
Other Eth Regular	43	-0.62	0.55	-0.63	0.06	26	-0.25	0.53	-0.22	0.08	2.72	15.73	0.001
Advanced 2 Parent	72	-0.35	0.70	-0.32	0.06	96	-0.25	0.49	-0.27	0.05	0.11	0.38	ns
Advanced Other Parent	25	-0.43	0.59	-0.50	0.09	32	-0.39	0.64	-0.33	0.08	0.40	2.13	ns
Regular 2 Parent	186	-0.45	0.67	-0.45	0.04	62	-0.31	0.54	-0.30	0.06	1.05	4.28	0.05
Regular Other Parent	130	-0.51	0.69	-0.53	0.05	47	-0.30	0.79	-0.26	0.08	2.41	7.53	0.01

Figure IX.4 Bar Graph of Adjusted Means from Table IX.4



As a final step in this series of ANCOVAs, we examined the relationship between a number of three-way combinations of these characteristics and change in test anxiety (Tables IX.5.a and IX.5.b, and Figures IX.5.a and IX.5.b). Four significant results emerge, one of which appears contrary to the expected effect of the TestEdge intervention. This is the result for Advanced White Males, which shows that this group of students in the control group achieved a greater reduction in test anxiety than their counterparts in the experimental group (-0.56 versus -0.24, p <0.05). Given the many thousands of statistical results generated in the analyses we conducted, and the anomalous nature of this finding, it is possible that this result is due to chance.

The other three significant results are in line with the expected effects of the intervention and are consistent with the findings presented above in relation to females and Regular Class status. All three results involve students in a Regular Class, and two of the three involve female students in a regular class. One group is White Females in a Regular Class (-0.71 versus -0.23, p <0.001), and the other group is Females in the Other ethnic category in a Regular Class (-0.72 versus -0.19, p <0.05). The third group is Other Males in a Regular Class (-0.56 versus -0.22, p <0.05).

Moving to the results in Table IX.5.b and Figure IX.5.b, two significant reductions in test anxiety are evident, both of which are consistent with the expected effects of the intervention. Again, both concern females in a Regular Class. The first is for Females in families with both biological parents (-0.55 versus -0.33, p <0.05), and the second is for Females in Other family situations (-0.69 versus -0.21, p <0.001).

Table IX.5.a ANCOVA of Post-Intervention Test Anxiety Change at Time 3 for Matched Sub-Samples on 3-way Combinations of Gender, Ethnicity, and Class Academic Level by Intervention Status

	Experimental Group					Control Group							
	N	Mean	SD	Adjusted Means	SEM	N	Mean	SD	Adjusted Means	SEM	Mean Sq.	F	*p* <
Regular White Males	50	-0.47	0.77	-0.45	0.08	34	-0.30	0.63	-0.34	0.09	0.228	0.79	ns
Regular Hispanic Males	92	-0.25	0.61	-0.26	0.05	7	-0.16	0.71	-0.03	0.19	0.342	1.425	ns
Regular Other Males	22	-0.45	0.49	-0.56	0.07	11	-0.43	0.59	-0.22	0.11	0.76	6.81	0.05
Advanced White Males	23	-0.25	0.50	-0.24	0.09	20	-0.54	0.63	-0.56	0.10	1.091	5.521	0.05
Advanced Hispanic Males	9	-0.44	0.55	-0.41	0.11	1	0.50	.	0.20	0.33	0.323	3.166	ns
Advanced Other Males	8	-0.25	0.73	-0.07	0.18	17	-0.32	0.57	-0.40	0.12	0.54	2.35	ns
Regular White Females	47	-0.72	0.72	-0.71	0.08	25	-0.21	0.80	-0.23	0.11	3.706	11.564	0.001
Regular Hispanic Females	82	-0.52	0.68	-0.54	0.06	16	-0.48	0.53	-0.36	0.13	0.414	1.467	ns
Regular Other Females	20	-0.78	0.56	-0.72	0.10	15	-0.11	0.46	-0.19	0.12	2.21	11.40	0.01
Advanced White Females	30	-0.35	0.61	-0.37	0.09	45	-0.18	0.47	-0.17	0.07	0.709	3.142	ns
Advance Hispanic Females	19	-0.78	0.64	-0.71	0.10	4	-0.09	0.31	-0.44	0.23	0.214	1.1	ns
Advanced Other Females	7	0.23	1.09	0.11	0.22	40	-0.33	0.52	-0.30	0.09	0.99	3.02	ns

ANCOVA

Figure IX.5.a Bar Graph of Adjusted Means from Table IX.5.a

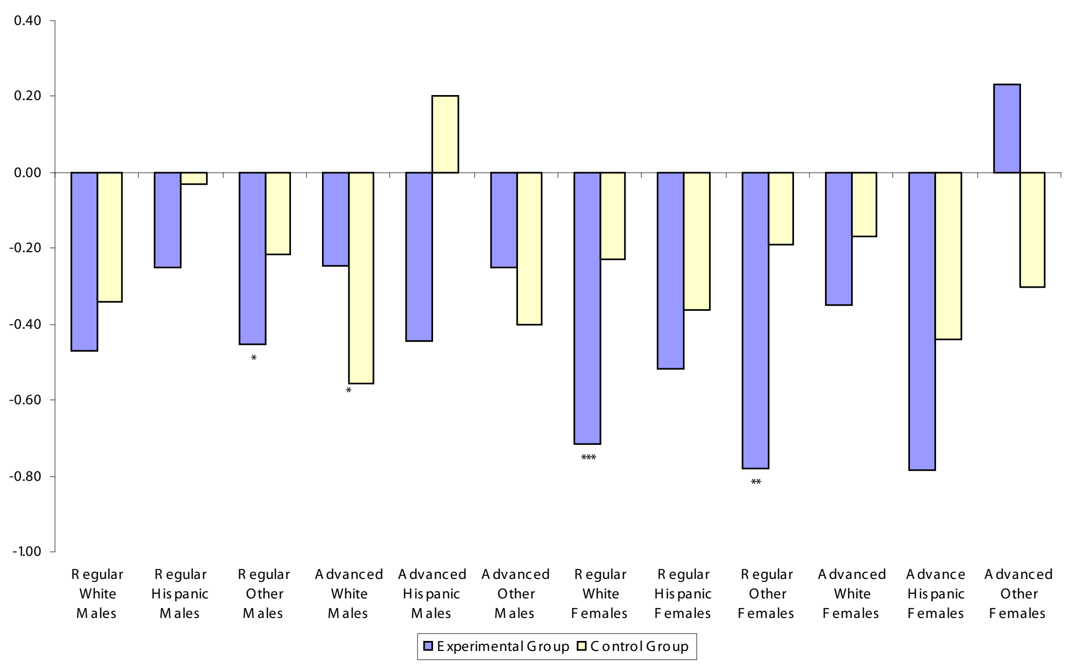
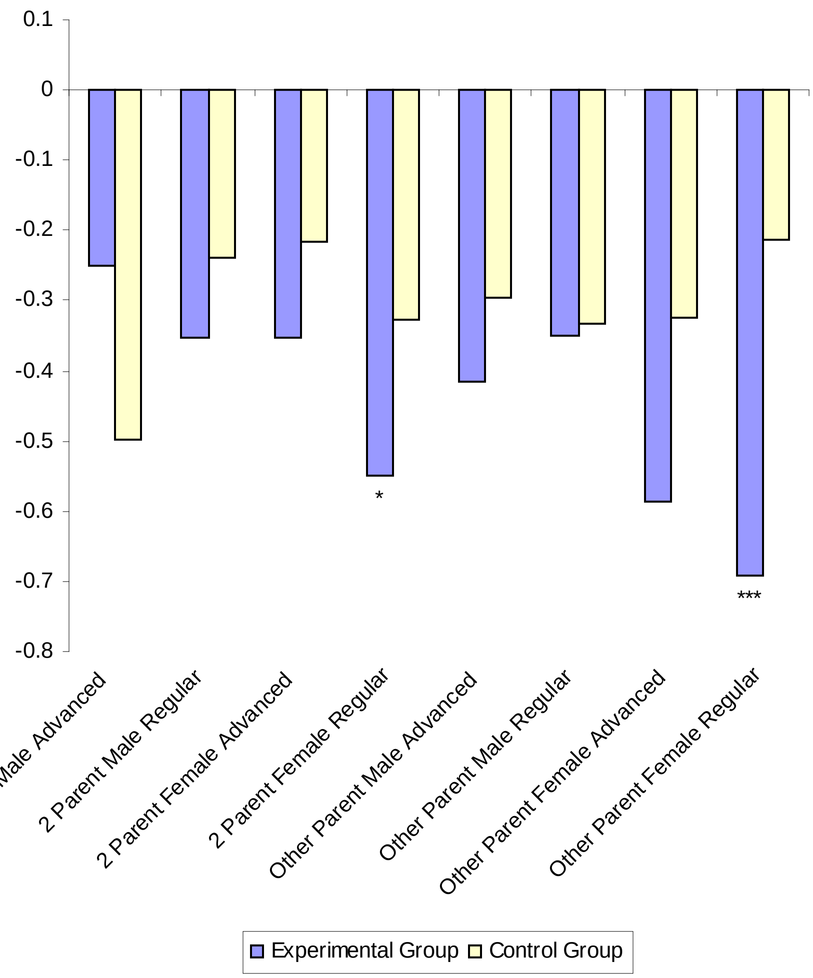


Table IX.5.b ANCOVA of Post-Intervention Test Anxiety Change at Time 3 for Matched Sub-Samples on 3-way Combinations of Gender, Other Parent Status, and Class Academic Level by Intervention Status

	Experimental Group					Control Group							
Test Anxiety delta (Time 3)	N	Mean	SD	Adjusted Mean	SEM adj.	N	Mean	SD	Adjusted Mean	SEM adj.	Mean Sq.	F	*p* <
2 Parent Male Advanced	33	-0.29	0.57	-0.25	0.08	25	-0.44	0.56	-0.50	0.10	0.85	3.83	ns
2 Parent Male Regular	96	-0.35	0.65	-0.35	0.05	32	-0.26	0.57	-0.24	0.08	0.31	1.51	ns
2 Parent Female Advanced	39	-0.39	0.80	-0.35	0.09	70	-0.19	0.45	-0.22	0.06	0.47	1.62	ns
2 Parent Female Regular	89	-0.55	0.68	-0.55	0.06	29	-0.33	0.44	-0.33	0.10	1.08	3.98	0.05
Other Parent Male Advanced	7	-0.29	0.50	-0.42	0.12	13	-0.37	0.74	-0.30	0.09	0.06	0.63	ns
Other Parent Male Regular	69	-0.33	0.66	-0.35	0.07	20	-0.40	0.71	-0.33	0.13	0.01	0.02	ns
Other Parent Female Advanced	18	-0.49	0.62	-0.59	0.11	19	-0.42	0.59	-0.32	0.11	0.60	2.92	ns
Other Parent Female Regular	61	-0.70	0.68	-0.69	0.07	27	-0.19	0.83	-0.21	0.10	4.29	14.64	0.001

ANCOVA

Figure IX.5.b Bar Graph of Adjusted Means from Table IX.5.b



Summary of Results

A strong pattern of results emerged from the analysis of the association between post intervention change in test anxiety and sub-samples matched on student characteristics:

- With only one exception, which appears an anomaly, every significant result has shown a consistent reduction in test anxiety in the experimental group sub-sample involved.
- Throughout the analysis of sub-samples matched on various combinations of student characteristics, two primary factors—female and regular class status—have consistently emerged as commonalities. A third factor which is present to a lesser degree is intact family status.

Change in Test Performance

Using the same ANCOVA approach, we turned our attention to student test performance when controlling for 9^{th} grade CST baseline differences. The baseline-adjusted mean test score on the 10^{th} grade CST ELA was used as the dependent variable and intervention status was used as the grouping variable. As in the analysis above, we started with category breakdowns for Gender, Ethnicity, Family Composition, and Class Academic Level as independent variables and then proceeded to investigate bivariate combinations of categories for these variables.

Starting with the results for the category breakdowns for Gender, Ethnicity,[13] Family Composition, and Class Academic Level (results not shown), there is only one significant result. This is the result for All Males, which shows that the males in the control group achieved higher test scores on average than their counterparts in the experimental group (354.80 versus 345.02, $p < 0.05$). For the bivariate combinations (results not shown), again there is only one significant result and it, too, involves males—males living in Other (non-intact) family situations. For these males in the control group, mean test scores were notably higher than those of their classmates in the experimental group (352.51 versus 334.78, $p < 0.01$). Finally, in relation to an ANCOVA of bivariate combinations of Class Level with Gender, Ethnicity, and Family Composition (results not shown), once again there is only one significant result, which also involves males.

[13] Not shown in Table IV.5 is the result of the comparison between the experimental group and control group for Asian students (387.14 versus 390.69, F = 0.283, $p = 0.597$; not significant).

Here, too, these males in the control group outperformed those in the experimental group (335.86 versus 325.11, p <0.05). While this set of findings runs counter to the expected effect of the TestEdge intervention, the consistent pattern of higher test performance of males in the control group suggests that it is likely a valid result.

To complete this series of analyses of bivariate combinations, we tightened the control on our dependent variable, the baseline measure of test performance, by holding 9th grade CST ELA test performance constant. This was achieved by constructing matched pairs of students (one from the experimental group and one from the control group) from the set of those who had identical 9th grade test scores at the beginning of the study. Of the thirty possible categories of bivariate combinations of student characteristics, eight categories were not available for this analysis due to small or non-existent cases when constructing the matched-pairs. For the remaining 22 categories, a matched-pairs t-test (2-tailed) difference in means procedure was performed, comparing the experimental and control groups on 10th grade CST ELA test performance.

Beginning with the bivariate combinations of the Gender, Ethnicity, and Family Composition categories (results not shown), no statistically significant results emerged. However, when these categories are combined with Academic Class Level (results not shown), three findings emerge which all involve students in Regular Classes for whom mean test performance is higher in the control group than in the experimental group. The first is for Males in a Regular Class (353.40 versus 334.00, t-test -3.10, p = 0.01)—a finding also revealed in the ANCOVA reported just above. The other two are new findings: one involves White students in a Regular Class (357.33 versus 345.21, t-test -2.16, p = 0.05); the other concerns students in a Regular Class with Two biological Parents (356.54 versus 342.64, t-test -2.59, p = 0.05). Although contrary to the expected effects of the intervention, the two consistent elements in these results (male and regular class status) suggests that they are likely valid.

Following the same logic as for the analysis of test anxiety, above, we then completed this series of ANCOVAs by examining the relationship between a number of three-way combinations and one four-way combination of these independent variables and pre-post change in test performance (results not shown). For the three-way combinations, two statistically significant results were found, both involving female students and both being consistent with the expected effects of the TestEdge intervention. The first result—White Females in a Regular Class—shows that these students in the experimental group outperformed their classmates in the control group on average by a margin of about 13 points (353.04 versus 339.18, $p < 0.05$). The other result is for Females in the Other ethnicity category in an Advanced Class: those in the experimental group outperformed their counterparts in the control group by a margin of almost 25 points (428.90 versus 405.52, $p < 0.01$).

There were no significant results for three-way combinations involving the categories of Family Composition with those of Gender and Class Academic Level. The one four-way combination examined—White Females living in a Two-Parent family in a Regular Class (data not shown)—did not produce in any significant results for pre-post change in test performance.

Summary of Results

When compared to the rich harvest of results on test anxiety and student characteristics above, fewer results of significance emerged from the analysis of the association between post intervention test performance and sub-samples matched on various sociodemographic characteristics. Two patterns are evident:

- Higher test performance by certain sub-samples in the control group over their counterparts in the experimental group. These sub-samples are consistently distinguished by two characteristics—male and regular class status.
- Higher test performance by two sub-samples in the experimental group—White Female Regular Class and Other Ethnicity Females in an Advanced Class. The common characteristic across these sub-samples is female.
- The common factor for the sub-samples of males and females for which higher test performance was observed is regular class status.

A summary of significant results from the analysis of the association of sub-samples matched on student characteristics to changes in test anxiety and test performance is presented in Table IX.6.

Table IX.6 Summary of Results: Association of Sub-Samples Matched on Student Characteristics with Post-Intervention Test Anxiety and Test Performance

Matched Sub-Samples Associated with:

Reduction in Test Anxiety, T1-T3	**Improvement in 9^{th}–10^{th} Grade CST ELA Test Score**
Univariate Grouping	
	Males: C>E*
Females: E<C***	
White: E<C**	
Other Ethnicity: E<C*	
Two Parent Family: E<C*	
Other Parent Family: E<C**	
Regular Class: E<C****	
Bivariate Grouping	
White Females: E<C***	
Females Two Parent Family: E<C**	
	Males Other Parent Family: C>E**
Females Other Parent Family: E<C**	
White Two Parent Family: E<C*	
Hispanic Two Parent Family: E<C*	
Other Ethnicity Two Parent Family: E<C*	
Other Ethnicity Other Parent Family: E<C***	
	Males Regular Class: C>E*
Females Regular Class: E<C****	
White Regular Class: E<C**	***White Regular Class: C>E****
Other Ethnicity Regular Class: E<C****	
Regular Class Two Parent Family: E<C*	***Regular Class Two Parent Family: C>E****
Regular Class Other Parent Family: E<C**	
Multivariate Grouping	
Males Other Ethnicity Regular Class: E<C*	
*Males White Advanced Class: C<E**	
Females White Regular Class: E<C***	**Females White Regular Class: E>C***
	Females Other Ethnicity Advanced Class: E>C**
Females Other Ethnicity Regular Class: E<C**	
Females Two Parent Regular Class: E<C*	
Females Other Parent Regular Class: E<C***	

KEY

E = Experimental School
C = Control School
Groupings in **bold** were selected for additional analysis
Findings in *italics* are not consistent with the expected effects of the TestEdge intervention.
Statistical significance: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$

Emotional Correlates of Test Anxiety and Test Performance Change

For the sub-samples of students just identified, the final step in our analysis was to investigate the relationship between various aspects of emotional disposition (as measured by the SOS scales) and the changes in test anxiety and test performance. Of the 24 groupings of students with different characteristics for which a significant reduction in test anxiety was observed, and the 7 groupings of students for which a significant increase in test performance was measured (see Table IX.6), we investigated 13 groupings[14] guided by the following considerations.

We first selected all groupings identified in the analysis of sub-samples above in which a reduction in test anxiety was associated with a change in test performance. There were three such groupings, all of which we investigated:

White Regular Class

Two Parent Regular Class

Female White Regular Class

We then examined the remaining 21 groupings in which only a reduction in test anxiety (i.e., without an associated improvement in test performance) was observed, and selected four groupings involving females for further investigation:

Female

Female Other Parent

Female Regular Class

Female Other Parent Regular Class

Next, we selected all groupings in which only an improvement in test performance (i.e., without a corresponding reduction in test anxiety) had occurred. There were four such groupings:

Male

Male Other Parent

Male Regular Class

Female Other Ethnicity Advanced Class

[14]A number of other groupings will also be analyzed in future work.

And finally, we selected two additional groupings involving Regular Class status in which a reduction in test anxiety had occurred:

Regular Class

Other Parent Regular Class

The results of these analyses are presented in Tables IX.7–IX.14.

1. Change in Test Anxiety Associated with Change in Test Performance

White Regular Class and Two Parent Regular Class. We begin with the three groupings in which both a change in test anxiety and a change in test performance were observed (Table IX.7). For two of these groupings—White Regular Class and Two Parent Regular Class, both used in the analysis of matched-pairs on ELA score above—the relationship between change in test anxiety and change in test performance is inconsistent with the expected effects of the intervention (as already noted above). Thus, while CST ELA mean score is higher in the control group for both groupings of students (344.98 versus 322.22, $p < 0.001$; and 344.08 versus 324.65, $p < 0.01$), test anxiety is significantly higher in the experimental group on all three measures of test anxiety (e.g., TAI-Global: 2.34 versus 1.89, $p < 0.001$; and 2.35 versus 1.93, $p < 0.001$). Interestingly, for both categories—White Regular Class and Two Parent Regular Class—Positive Affect is also higher in the control group than it is in the experimental group (3.01 versus 2.75, $p < 0.01$; and 2.96 versus 2.74, $p < 0.01$). These relationships between reduced test anxiety, increased positive affect, and increased test performance in the control group are consistent with the emotion-based theory of test anxiety, presented at the beginning of this report.[15]

Returning to the results, although Emotional Discord is higher in the control group of the Two Parent Regular Class grouping (1.97 versus 1.82, $p < 0.05$), because this result is close to marginal significance and is also inconsistent with the other results of reduced test anxiety and increased positive affect, it is likely an anomalous result produced by chance.

[15]It is possible that some other unmeasured factor/s—such as higher socioeconomic status, better student academic ability or motivation, etc., in these control group sub-samples—and/or deficient implementation of the TestEdge program in these experimental group sub-samples could produce this outcome.

Table IX.7 Association of Post-Intervention Test Anxiety, Test Performance, and Measures of Emotional Disposition with Intervention Status in Sub-Samples Matched on Selected Student Characteristics

	Experimental Group				Control Group						
	N	Mean	SD	SEM	N	Mean	SD	SEM	t	df	*p* <
White Regular Classes											
CST English-Language Arts 10	41	322.22	47.02	7.34	41	344.98	39.53	6.17	3.63	40	0.001
Test Anxiety-Global	41	2.34	0.85	0.13	41	1.89	0.75	0.12	4.86	40	0.001
Test Anxiety-Worry	41	2.37	0.90	0.14	41	2.02	0.79	0.12	2.88	40	0.01
Test Anxiety-Emotional	41	2.32	0.91	0.14	41	1.76	0.79	0.12	5.14	40	0.001
Teacher Support	43	2.98	0.84	0.13	43	3.15	0.71	0.11	-1.78	42	ns
Parental Support	41	3.51	0.71	0.11	41	3.68	0.55	0.09	-1.90	40	ns
Positive Affect	43	2.75	0.60	0.09	43	3.01	0.61	0.09	-3.35	42	0.01
Emotional Discord	43	1.96	0.71	0.11	43	2.03	0.76	0.12	-0.81	42	ns
2 Parent Regular Classes											
CST English-Language Arts 10	37	324.65	49.48	8.14	37	344.08	43.94	7.22	3.06	36	0.01
Test Anxiety-Global	39	2.35	0.87	0.14	39	1.93	0.79	0.13	4.41	38	0.001
Test Anxiety-Worry	39	2.34	0.90	0.14	39	2.02	0.85	0.14	2.75	38	0.01
Test Anxiety-Emotional	39	2.37	0.95	0.15	39	1.84	0.87	0.14	4.38	38	0.001
Teacher Support	39	2.83	0.93	0.15	39	2.99	0.75	0.12	-1.46	38	ns
Parental Support	37	3.51	0.71	0.12	37	3.53	0.67	0.11	-0.22	36	ns
Positive Affect	39	2.74	0.64	0.10	39	2.96	0.61	0.10	-2.88	38	0.01
Emotional Discord	39	1.82	0.57	0.09	39	1.97	0.65	0.10	-2.07	38	0.05

Paired t-Test

White Females - Regular classes

	Experimental Group					Control Group							
Dependent	N	Mean	SD	Adjusted Mean	SEM	N	Mean	SD	Adjusted Mean	SEM	Mean Sq	F	*p* <
CST English-Language Arts 10	49	344.39	29.21	353.04	3.67	27	354.89	53.73	339.18	5.04	2914.39	4.64	0.05
Test Anxiety-Delta (Time 3)	47	-0.72	0.72	-0.71	0.08	25	-0.21	0.80	-0.23	0.11	3.71	11.56	0.001
Test Anxiety-Global	51	2.24	0.93	2.21	0.09	29	2.47	0.81	2.52	0.12	1.79	4.33	0.05
Test Anxiety-Worry	51	2.34	0.97	2.29	0.10	29	2.54	0.80	2.64	0.13	2.21	4.43	0.05
Test Anxiety-Emotional	51	2.14	0.96	2.14	0.10	29	2.40	0.97	2.40	0.14	1.29	2.41	ns
Teacher Support	53	3.15	0.79	3.09	0.09	29	2.70	0.82	2.81	0.12	1.47	3.56	ns
Parental Support	50	3.49	0.68	3.44	0.08	28	3.08	0.82	3.18	0.11	1.12	3.61	ns
Positive Affect	53	2.87	0.70	2.82	0.08	29	2.47	0.65	2.56	0.10	1.16	3.78	ns
Emotional Discord	53	2.46	0.81	2.51	0.09	29	2.66	0.70	2.59	0.12	0.12	0.28	ns

ANCOVA

White Female Regular Class. Turning to the results (the bottom of Table IX.7)—White Female Regular Class—this is the one grouping of students for which both the change in test anxiety and change in test performance were consistent with the effects expected from the intervention. Thus, we see that higher test performance is associated with lower test anxiety in the experimental group while the obverse holds for the control group (ELA mean score: 353.04 versus 339.18, $p < 0.05$; TAI-Global: 2.21 versus 2.52, $p < 0.05$). There is some evidence that these changes in test anxiety and test performance may be related to perceived changes in social support (Teacher Support, 3.09 versus 2.81, $p = 0.06$; Parental Support, 3.44 versus 3.18, $p = 0.06$) and emotional state (Positive Affect, 2.82 versus 2.56, $p = 0.06$). Even though these results are all marginally significant, the pattern across the three measures is consistent, which suggests that these results are likely to be valid.

A picture of the changes in test anxiety and test performance for students in this category on a case-by-case basis is presented in Figure IX.8.

Figure IX.8.a Bar Graph Showing 9th–10th Grade CST ELA Score Change (descending order) by Student in the White Female Regular Class Matched-Group Sub-Sample by Intervention Status

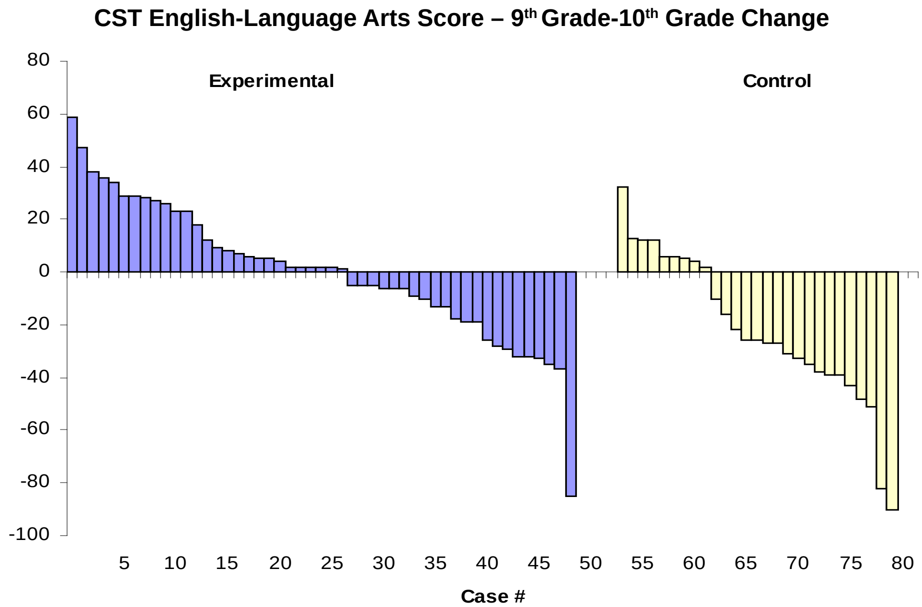
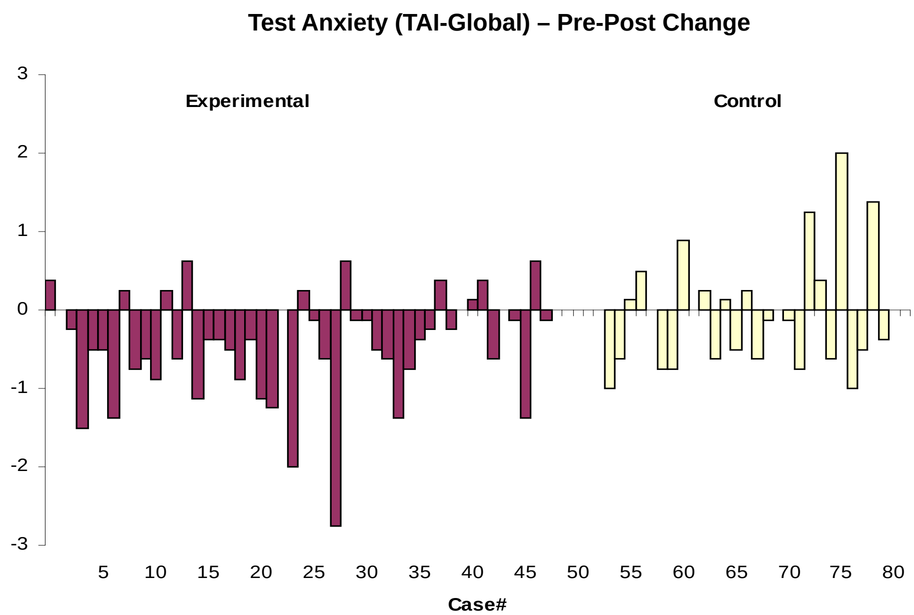


Figure IX.8.b Bar Graph Showing Pre-Post Change in Test Anxiety by Student (presented in descending order on 9th–10th grade CST ELA Score Change) in the White Female Regular Class Matched-Group Sub-Sample by Intervention Status



The top bar graph shows the 9th–10th grade CST ELA score change for each student (by descending order) in the White Females Regular Class category grouped by intervention status. Not only did a greater proportion of students improve their test performance in the experimental group (55.1% versus 33.3% for the control group), but the improvement in their test scores was typically of a greater magnitude. The obverse pattern can be seen for a decrease in test performance. The bottom bar graph shows the pre to post change in test anxiety for each student, presented also in descending order of the 9th–10th grade CST ELA score change. A difference in the pattern of test anxiety change is apparent in the two groups. Whereas 80% of those in the experimental group saw a reduction in their test anxiety, almost half as many (37%) in the control group reported an increase in test anxiety by the end of the study.

Table IX.9 shows that the relationship between the incidence of reduced test anxiety and increased CST ELA test performance was marginally significant (Pearson $Chi^2 = 3.63$, $p = 0.06$). Given the weight of evidence from other findings for this group of students in the experimental group, we are inclined to accept this result as valid.

Table IX.9 Contingency Table Analysis of Relationship Between Test Anxiety and CST ELA Change in the White Female Regular Class Matched-Group Sub-Sample by Intervention Status

	Decreased TAI Decreased ELA	Decreased TAI Increased ELA	Total no. of cases
Experimental	14	20	34
Control	10	4	14
Total	24	24	48
X^2 =	3.63		
***p* =**	0.057		
df	1		

	Increased TAI Decreased ELA	Increased TAI Increased ELA	Total no. of cases
Experimental	7	6	13
Control	8	5	13
Total	15	11	26
X^2 =	0.16		
***p* =**	0.691		
df	1		

2. Reduction in Test Anxiety or Improvement in Test Performance

For this analysis of the four groupings in which either a reduction in test anxiety or an increase in test performance was observed (Table IX.10), we added the complementary category for purposes of comparison—viz, in presenting the results for Females we also added the results for Males on the same SOS scales.

Female. Beginning with the results for Females, a significant reduction in test anxiety is evident in the experimental group, when compared to the control group, across all four measures of test anxiety (Test Anxiety-Delta, Time 3: -0.54 versus -0.28; TAI-Global: 2.20 versus 2.55; TAI-Worry: 2.22 versus 2.51; TAI-Emotionality: 2.18 versus 2.59; *p* for all <0.001). There was also a greater level Positive Affect (2.84 versus 2.74, *p* <0.05), which is consistent with the observed reduction in test anxiety.

Male. Although the Males in the control group outperformed their counterparts in the experimental group on the CST ELA (354.80 versus 345.02, *p* <0.05), there is also some evidence that they experienced an increase in test anxiety (TAI-Emotional: 1.98 versus 1.78, *p* <0.05). However, because this is an isolated result and close to marginal significance, there is a good likelihood that it is the result of chance.

Table IX.10 Association of Post-Intervention Test Anxiety, Test Performance, and Measures of Emotional Disposition with Intervention Status in Sub-Samples Matched on Gender

	Experimental group					Control group							
Dependent	N	Mean	SD	Adjusted Mean	SEM	N	Mean	SD	Adjusted Mean	SEM	Mean Sq	F	*p* <
Females													
CST English - Language Arts 10	222	351.86	49.48	364.95	1.91	148	382.83	57.18	363.19	2.37	244.49	0.32	ns
Test Anxiety-Delta (Time 3)	207	-0.56	0.70	-0.54	0.04	145	-0.25	0.56	-0.28	0.04	5.80	20.75	0.001
Test Anxiety-Global	231	2.24	0.83	2.20	0.04	150	2.49	0.87	2.55	0.05	10.79	28.80	0.001
Test Anxiety-Worry	230	2.29	0.88	2.22	0.04	150	2.41	0.90	2.51	0.05	7.35	17.71	0.001
Test Anxiety-Emotional	229	2.20	0.90	2.18	0.05	144	2.57	0.98	2.59	0.06	14.69	27.31	0.001
Positive Class Experience	228	3.01	0.72	2.95	0.04	147	2.73	0.68	2.83	0.05	1.10	3.63	ns
Positive Affect	235	2.86	0.66	2.84	0.03	154	2.70	0.63	2.74	0.04	0.93	4.01	0.05
Emotional Discord	235	2.28	0.78	2.30	0.04	153	2.43	0.81	2.41	0.05	1.19	3.43	ns
Stress Management	233	2.41	0.64	2.37	0.03	154	2.41	0.64	2.47	0.04	0.89	3.46	ns
Males													
CST English - Language Arts 10	227	338.76	51.55	345.02	2.10	97	369.44	55.51	354.80	3.25	6188.63	6.25	0.05
Test Anxiety-Delta (Time 3)	205	-0.33	0.63	-0.33	0.03	90	-0.36	0.62	-0.35	0.05	0.03	0.13	ns
Test Anxiety-Global	244	1.85	0.80	1.85	0.04	99	1.94	0.85	1.95	0.06	0.79	2.42	ns
Test Anxiety-Worry	244	1.94	0.86	1.91	0.04	99	1.87	0.82	1.93	0.06	0.02	0.06	ns
Test Anxiety-Emotional	241	1.77	0.82	1.78	0.04	99	2.01	0.98	1.98	0.07	2.65	6.40	0.05
Positive Class Experience	232	2.91	0.74	2.93	0.04	93	2.87	0.78	2.83	0.07	0.64	1.49	ns
Positive Affect	250	2.82	0.65	2.82	0.03	101	2.79	0.68	2.79	0.05	0.08	0.28	ns
Emotional Discord	247	1.97	0.70	1.97	0.04	101	2.02	0.69	2.01	0.06	0.13	0.37	ns
Stress Management	248	2.42	0.67	2.42	0.04	101	2.41	0.63	2.42	0.06	0.00	0.01	ns

ANCOVA

Female Other Parents. A significant reduction in test anxiety is evident across all four measures of test anxiety for Females Other Parents (Table IX.11) in the experimental group, when compared to the control group (Test Anxiety-Delta, Time 3: -0.66 versus -0.27; TAI-Global: 2.05 versus 2.64; TAI-Worry: 2.10 versus 2.72; TAI-Emotional: 2.00 versus 2.58; *p* for all <0.001). There was a greater level of Positive Affect as well (2.81 versus 2.63, *p* <0.05), which is consistent with the observed decline in test anxiety.

Male Other Parents. Although Males with Other Parents in the control group outperformed their classmates in the experimental group on the CST ELA (351.95 versus 334.03, *p* <0.01), there were no significant results for test anxiety or any measures of emotional disposition.

Table IX.11 Association of Post-Intervention Test Anxiety, Test Performance, and Measures of Emotional Disposition with Intervention Status in Sub-Samples Matched on Gender and Other Parents Family Status

	Experimental group					Control group							
Dependent	N	Mean	SD	Adjusted Mean	SEM	N	Mean	SD	Adjusted Mean	SEM	Mean Sq	F	*p* <
Females - Other parents													
CST English - Language Arts 10	81	347.48	47.74	355.89	3.38	48	366.10	62.46	351.91	4.42	450.22	0.50	ns
Test Anxiety-Delta (Time 3)	79	-0.65	0.67	-0.66	0.06	46	-0.28	0.74	-0.27	0.08	4.54	17.02	0.001
Test Anxiety-Global	87	2.07	0.82	2.05	0.07	47	2.60	0.96	2.64	0.09	10.79	26.72	0.001
Test Anxiety-Worry	87	2.15	0.88	2.10	0.07	47	2.62	0.99	2.72	0.10	11.63	26.10	0.001
Test Anxiety-Emotional	85	1.98	0.85	2.00	0.08	46	2.62	1.03	2.58	0.11	10.20	19.33	0.001
Positive Affect	89	2.85	0.59	2.81	0.05	49	2.56	0.58	2.63	0.07	1.01	4.42	0.05
Males - Other Parents													
CST English - Language Arts 10	83	326.55	47.94	334.03	3.44	37	368.73	53.97	351.95	5.24	7586.85	7.91	0.01
Test Anxiety-Delta (Time 3)	76	-0.33	0.64	-0.34	0.06	33	-0.39	0.71	-0.36	0.09	0.00	0.02	ns
Test Anxiety-Global	91	1.90	0.86	1.91	0.06	37	1.95	0.73	1.92	0.10	0.00	0.00	ns
Test Anxiety-Worry	91	1.96	0.90	1.95	0.07	37	1.84	0.72	1.85	0.11	0.29	0.70	ns
Test Anxiety-Emotional	91	1.84	0.89	1.87	0.07	37	2.05	0.84	1.98	0.11	0.36	0.81	ns
Positive Affect	93	2.82	0.56	2.81	0.05	39	2.63	0.66	2.64	0.08	0.77	2.97	ns

ANCOVA

Female Regular Class. A more extensive set of findings emerges for Females in a Regular Class (Table IX.12). Again, a significant reduction in test anxiety is evident across all four measures of test anxiety for females in this category in the experimental group, compared to those in the control group (Test Anxiety-Delta, Time 3: -0.61 versus -0.27; TAI-Global: 2.27 versus 2.68; TAI-Worry: 2.35 versus 2.74; TAI-Emotional: 2.18 versus 2.64; *p* for all <0.001). There were other significant changes in the SOS scales observed as well: greater Extent of Friendship (3.31 versus 3.17, *p* <0.05) and greater Positive Affect (2.76 versus 2.58, *p* <0.05), accompanied by lower Negative Affect (2.32 versus 2.50, *p* <0.05), lower Emotional Discord (2.41 versus 2.63, *p* <0.05), and lower Interactional Difficulty (1.96 versus 2.16, *p* <0.05). Although close to marginal significance, the consistency in this pattern of results on the SOS scales suggests they are likely to be valid.

Male Regular Class. By contrast, while the Males in the control group outperformed their counterparts in the experimental group on the CST ELA (335.86 versus 325.11, *p* <0.05), this was associated with an increase in test anxiety (TAI-Emotional scale: 2.02 versus 1.75, *p* <0.01). This result for the males is consistent with a pattern already documented above.

Table IX.12 Association of Post-Intervention Test Anxiety, Test Performance, and Measures of Emotional Disposition with Intervention Status in Sub-Samples Matched on Gender and Regular Class Status

	Experimental group					Control group							
Dependent	N	Mean	SD	Adjusted Mean	SEM	N	Mean	SD	Adjusted Mean	SEM	Mean Sq	F	*p* <
Females - Regular Classes													
CST English - Language Arts 10	157	330.75	36.83	333.75	1.97	56	338.46	53.33	330.05	3.32	555.30	0.91	ns
Test Anxiety-Delta (Time 3)	150	-0.61	0.68	-0.61	0.04	56	-0.26	0.65	-0.27	0.07	4.62	16.19	0.001
Test Anxiety-Global	164	2.28	0.83	2.27	0.05	58	2.64	0.87	2.68	0.09	7.35	17.46	0.001
Test Anxiety-Worry	163	2.37	0.87	2.35	0.05	58	2.67	0.89	2.74	0.09	6.46	14.16	0.001
Test Anxiety-Emotional	163	2.18	0.89	2.18	0.06	55	2.63	0.95	2.64	0.10	8.72	15.49	0.001
Extent of Friendship	166	3.31	0.53	3.31	0.04	60	3.16	0.64	3.17	0.06	0.91	4.13	0.05
Positive Affect	167	2.80	0.66	2.76	0.04	60	2.49	0.68	2.58	0.06	1.41	5.87	0.05
Negative Affect	167	2.28	0.72	2.32	0.04	60	2.61	0.77	2.50	0.08	1.46	4.42	0.05
Emotional Discord	167	2.38	0.77	2.41	0.05	59	2.72	0.78	2.63	0.08	1.98	5.35	0.05
Interactional Difficulty	167	1.94	0.64	1.96	0.04	60	2.23	0.68	2.16	0.07	1.68	5.90	0.05
Males - Regular Classes													
CST English - Language Arts 10	183	324.23	44.44	325.11	2.27	52	338.94	47.72	335.86	4.25	4666.80	4.97	0.05
Test Anxiety-Delta (Time 3)	165	-0.34	0.65	-0.35	0.04	52	-0.31	0.63	-0.28	0.07	0.24	0.95	ns
Test Anxiety-Global	200	1.84	0.81	1.85	0.04	54	2.05	0.85	2.01	0.08	1.01	2.92	ns
Test Anxiety-Worry	200	1.95	0.88	1.95	0.05	54	2.00	0.83	2.00	0.09	0.07	0.17	ns
Test Anxiety-Emotional	199	1.73	0.83	1.75	0.05	54	2.10	0.96	2.02	0.09	3.09	7.42	0.01
Extent of Friendship	205	2.59	0.72	2.62	0.04	54	2.71	0.63	2.62	0.08	0.00	0.01	ns
Positive Affect	205	2.81	0.66	2.81	0.04	55	2.75	0.66	2.76	0.07	0.08	0.30	ns
Negative Affect	205	1.93	0.64	1.94	0.04	55	2.03	0.71	2.01	0.07	0.25	1.00	ns
Emotional Discord	202	1.97	0.74	1.98	0.04	55	2.01	0.65	2.00	0.08	0.02	0.07	ns
Interactional Difficulty	205	1.87	0.63	1.89	0.03	55	2.00	0.58	1.95	0.07	0.17	0.72	ns

ANCOVA

Female Other Ethnicity Advanced Class. Finally, there are some findings of note for Females of Other ethnic origin in an Advanced Class (Table IX.13). Although the females in the experimental group achieved a higher (baseline-adjusted) mean ELA 10th grade score than those in the control group (428.90 versus 405.52, *p* <0.05), they had less positive Feelings About School (3.48 versus 3.83, *p* <0.05), were somewhat less positive about their Educational Plans (3.57 versus 3.78, *p* <0.05), and were also less positive about Life Preparedness (2.53 versus 2.94, *p* <0.05). When these results are combined with the lack of an expected reduction in test anxiety, they suggest that the observed increase in test performance may be associated with factors other than the TestEdge intervention. There was only one significant result for Males of Other ethnic origin in an Advanced Class—Parental Support (3.76 versus 3.49, p <0.05)—which was also higher in the control group. However, given the very small case counts in the experimental group for the female (N = 9) and male (N = 7-8) sub-samples of each of these categories, it is likely that these results are not statistically stable; caution must therefore be exercised in interpreting them.

Table IX.13 Association of Post-Intervention Test Anxiety, Test Performance, and Measures of Emotional Disposition with Intervention Status in Sub-Samples Matched on Gender/Other Ethnicity/Advanced Class Status

	Experimental group					Control group							
Dependent	N	Mean	SD	Adjusted Mean	SEM	N	Mean	SD	Adjusted Mean	SEM	Mean Sq	F	p <
Females - Other ethnicity - Advanced classes													
CST English - Language Arts 10	9	430.33	35.84	428.90	8.67	40	405.20	31.86	405.52	4.11	4008.14	5.94	0.05
Feelings about School	9	3.11	0.88	3.48	0.15	41	3.91	0.51	3.83	0.07	0.79	4.74	0.05
Teacher Support	9	3.00	0.94	2.92	0.17	41	3.26	0.66	3.28	0.08	0.93	3.77	ns
Educational Plans	9	3.59	1.00	3.57	0.09	41	3.77	0.29	3.78	0.04	0.33	4.60	0.05
Life Preparedness	9	2.48	0.93	2.53	0.15	41	2.95	0.64	2.94	0.07	1.26	6.08	0.05
Parental Support	9	3.68	0.34	3.69	0.13	38	3.54	0.68	3.54	0.06	0.16	1.14	ns
Interactional Difficulty	9	1.64	0.55	1.60	0.13	41	1.87	0.53	1.88	0.06	0.61	4.33	0.05
Males - Other ethnicity - Advanced classes													
CST English - Language Arts 10	8	395.50	40.03	386.00	8.69	19	384.32	36.66	388.32	5.58	28.73	0.05	ns
Feelings about School	8	3.58	0.56	3.71	0.16	20	3.68	0.56	3.63	0.10	0.03	0.15	ns
Teacher Support	8	2.96	0.95	2.85	0.22	20	3.13	0.72	3.18	0.14	0.59	1.62	ns
Educational Plans	8	3.67	0.40	3.68	0.09	20	3.90	0.24	3.90	0.06	0.27	3.82	ns
Life Preparedness	8	3.00	0.73	3.11	0.21	20	3.22	0.68	3.17	0.13	0.02	0.06	ns
Parental Support	8	3.56	0.46	3.49	0.09	20	3.73	0.39	3.76	0.06	0.41	6.79	0.05
Interactional Difficulty	7	2.20	0.66	2.11	0.23	20	1.95	0.60	1.98	0.13	0.08	0.22	ns

ANCOVA

3. Reduction in Test Anxiety for Certain Characteristics and Regular Class Status

Moving to the final three groupings of students (see Table IX.14)—those for whom there was a significant reduction in test anxiety for certain characteristics combined with membership in a Regular Class (Regular Class, Other Parent Regular Class, Female Other Parent Regular Class)—a number of emotional correlates are evident for the first two categories.

Regular Class. The strongest pattern of test anxiety reduction linked with other emotional changes can be seen for students in the experimental group in a Regular Class.

For this grouping of students, the strong decline in test anxiety (TAI-Global: 2.04 versus 2.33, *p* <0.001) was accompanied by a significant increase in Positive Affect (2.79 versus 2.67, *p* <0.05) and a corresponding decrease in Negative Affect (2.11 versus 2.26, *p* <0.05), Emotional Discord (1.93 versus 2.06, *p* <0.05), and Interactional Difficulty (2.37 versus 2.48, *p* <0.05). Given their consistency, it is likely that these SOS results are valid. Since the majority of students in high school are in regular classes at a non-advanced academic level, these findings are compelling evidence of the impact of the TestEdge intervention on effecting positive changes in the emotional states of a large portion of the 10th grade student population; for the Regular Class sub-sample in the experimental group, in Table IX.14, the number of cases (N) range from 318-375.[16]

Table IX.14 Association of Post-Intervention Test Anxiety, Test Performance, and Measures of Emotional Disposition with Intervention Status in Sub-Samples Matched on Regular Class Status and Other Selected Characteristics

	Experimental group					Control group							
Dependent	N	Mean	SD	Adjusted Mean	SEM	N	Mean	SD	Adjusted Mean	SEM	Mean Sq	F	*p* <
Regular Classes													
Test Anxiety-Delta (Time 3)	318	-0.47	0.68	-0.48	0.03	112	-0.32	0.65	-0.29	0.05	3.17	11.61	0.001
Test Anxiety-Global	366	2.03	0.85	2.04	0.03	116	2.36	0.90	2.33	0.06	7.13	18.26	0.001
Test Anxiety-Worry	365	2.14	0.90	2.13	0.04	116	2.33	0.91	2.34	0.06	3.81	8.32	0.01
Test Anxiety-Emotional	364	1.93	0.88	1.95	0.04	113	2.38	0.98	2.32	0.07	11.89	24.22	0.001
Positive Affect	375	2.80	0.66	2.79	0.03	119	2.63	0.69	2.67	0.05	1.14	4.39	0.05
Negative Affect	375	2.09	0.70	2.11	0.03	119	2.33	0.79	2.26	0.05	1.87	6.29	0.05
Emotional Discord	372	2.16	0.79	2.18	0.03	118	2.37	0.80	2.31	0.06	1.50	3.95	0.05
Interactional Difficulty	375	1.91	0.64	1.93	0.03	119	2.12	0.65	2.06	0.05	1.49	5.69	0.05
Stress Management	373	2.40	0.68	2.37	0.03	119	2.41	0.68	2.48	0.05	1.07	3.47	ns
Other Parent - Regular Classes													
Test Anxiety-Delta (Time 3)	132	-0.51	0.69	-0.53	0.05	50	-0.32	0.78	-0.27	0.08	2.38	7.64	0.01
Test Anxiety-Global	153	2.01	0.85	2.02	0.05	51	2.35	0.88	2.31	0.09	3.22	7.69	0.01
Test Anxiety-Worry	153	2.10	0.90	2.10	0.06	51	2.36	0.91	2.37	0.10	2.78	5.56	0.05
Test Anxiety-Emotional	152	1.90	0.86	1.93	0.06	50	2.36	0.95	2.26	0.10	4.11	8.41	0.01
Positive Affect	156	2.83	0.58	2.79	0.04	53	2.53	0.65	2.63	0.07	1.03	4.28	0.05
Negative Affect	156	2.12	0.70	2.16	0.05	53	2.48	0.71	2.37	0.08	1.61	4.76	0.05
Emotional Discord	154	2.23	0.79	2.27	0.05	52	2.52	0.73	2.38	0.09	0.47	1.24	ns
Interactional Difficulty	156	1.94	0.64	1.97	0.04	53	2.18	0.62	2.09	0.07	0.55	1.94	ns
Stress Management	156	2.39	0.65	2.35	0.05	53	2.33	0.63	2.46	0.08	0.48	1.54	ns
Females - Other Parent - Regular Classes													
Test Anxiety-Delta (Time 3)	61	-0.70	0.68	-0.69	0.07	27	-0.19	0.83	-0.21	0.10	4.29	14.64	0.001
Test Anxiety-Global	68	2.13	0.80	2.10	0.08	28	2.57	0.94	2.65	0.13	5.77	13.35	0.001
Test Anxiety-Worry	68	2.26	0.86	2.21	0.09	28	2.69	0.95	2.82	0.13	7.28	14.67	0.001
Test Anxiety-Emotional	67	1.98	0.82	1.99	0.09	27	2.51	1.01	2.49	0.14	4.79	9.46	0.01
Positive Affect	69	2.82	0.61	2.76	0.06	29	2.49	0.59	2.64	0.09	0.28	1.31	ns
Negative Affect	69	2.27	0.71	2.32	0.07	29	2.65	0.75	2.52	0.12	0.73	1.94	ns
Emotional Discord	69	2.39	0.77	2.43	0.08	28	2.75	0.76	2.65	0.12	0.94	2.43	ns
Interactional Difficulty	69	1.96	0.60	1.99	0.07	29	2.26	0.62	2.18	0.10	0.71	2.43	ns
Stress Management	69	2.36	0.67	2.30	0.07	29	2.22	0.59	2.38	0.12	0.11	0.31	ns

ANCOVA

[16] The variance in N is due to list-wise deletion of "missing cases" in the ANCOVA procedure.

Other Parent Regular Class. For students in the Other Parent Regular Class grouping, there were fewer emotional correlates associated with the reduction in test anxiety. Thus while test anxiety is significantly lower in the experimental group (e.g., TAI-Global: 2.02 versus 2.31, $p < 0.01$), it is associated with an increase in Positive Affect (2.79 versus 2.63, $p < 0.05$) and a corresponding decrease in Negative Affect (2.16 versus 2.37, $p < 0.05$). Although the latter are close to marginal statistical significance, they are consistent with each other and also consistent with the reduction in test anxiety and, therefore, likely valid.

Female Other Parent Regular Class. For the last category examined—Female with Other Parents in a Regular Class—the significant decline in test anxiety observed in the experimental group (TAI-Global: 2.10 versus 2.65, $p < 0.001$) was not associated with any significant improvement in emotional disposition.

Summary of Results

Across the 13 sub-samples constructed from different combinations of student characteristics, we found a number of commonalities in emotional disposition that appear associated with the changes in test anxiety and test performance:

- A reduction in test anxiety and an improvement in test performance was likely to be associated with an increase in positive affect, in both the experimental and control groups.
- There is evidence for females in the experimental group that changes in test anxiety and test performance are also associated with increased social support from teachers and parents.
- A reduction in test anxiety without a corresponding increase in test performance was highly likely to be associated with an increase in positive affect.
- Such a reduction in test anxiety was often also associated with a reduction in negative affect, emotional discord, and interactional difficulty.

Sub-Samples Matched on Class Test Performance

An alternative to constructing matched-group comparisons on the sociodemographic and class academic level variables, just presented, is to use the 9th grade baseline test performance of students as the matching or controlling variable. Any observed differences in test performance on the 10th grade test would then be treated as the dependent variable in follow-up analyses conducted to explain the difference by using the sociodemographic variables and scales from the SOS as independent variables.

One way to control for differences in test performance is to construct matched-pairs of individual students (one from the experimental group and one from the control group) whose test scores on a given test are the same. However, because separate analyses (not shown) of matched pairs of students constructed from their 9th grade CST scores in Mathematics and English-Language Arts, respectively, failed to yield any definitive results, we pursued a second approach.

Construction of Matched Groups Using Baseline CST Scores

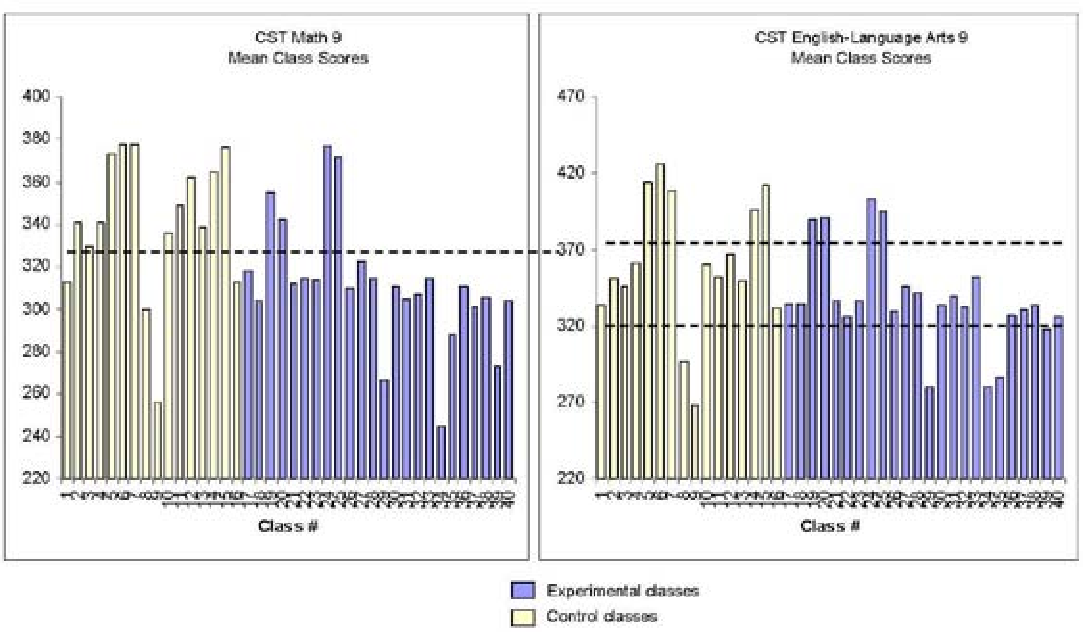
For the second approach, matched-groups of classes were constructed by matching classes by mean class test performance on the 9th grade CST. Separate analyses using 9th grade Mathematics test scores and English-Language Arts test scores, respectively, were conducted to identify any 10th grade test performance difference, and then to attempt to explain the observed difference with the sociodemographic variables and SOS scales.

For the first analysis, two criteria were used to construct the matched-class Mathematics sub-samples. The first was the 9th grade mean CST score for each class. This was computed for the two CST tests and displayed in separate bar charts (one for Mathematics and one for ELA, as shown in Figure IX.15), which plot the class mean test scores by intervention status. On the basis of the distribution of class means, three possible groupings of matched-class comparisons were attempted and achieved for the 9th grade ELA score class distribution: high-scoring classes, mid-scoring classes, and low-scoring classes. Because various combinations of Mathematics tests were taken by different groups of students, as previously noted, a second criterion had to be added for constructing the matched-class comparisons on Mathematics. This criterion included students who took either one of the two different combinations of 9th and 10th grade geometry and algebra classes. This operational logic created the possibility of mutually exclusive groups matched by mathematical subject matter and by performance level. Because there were no students who met this second criterion in two of the classes in the control group, these classes were excluded from the analysis that follows.

Table IX.15 Construction of Math Groups 1 and 2: ANCOVA Results for 10th Grade CST Mathematics and English-Language Arts by Intervention Status

	Experimental Group					Control Group							
	N	Mean	SD	Adjusted Mean	SEM	N	Mean	SD	Adjusted Mean	SEM	Mean Sq.	F	*p* <
Math Group 1 (Geometry & Algebra 2)													
Math high-score classes													
CST Math 10	68	359.10	53.12	359.71	4.36	60	361.27	40.76	360.58	4.64	24.60	0.02	ns
CST English-Language Arts 10	69	412.13	33.29	413.44	3.35	60	404.47	37.67	402.96	3.60	3505.71	4.52	0.05
ELA high-score classes													
CST Math 10	68	359.10	53.12	360.00	4.40	52	363.75	38.90	362.57	5.03	193.85	0.15	ns
CST English-Language Arts 10	69	412.13	33.29	413.23	3.25	52	406.10	35.96	404.63	3.74	2186.38	3.01	ns
ELA mid-score classes													
CST Math 10	48	316.23	38.64	318.12	4.34	9	343.11	48.40	333.03	10.12	1640.51	1.82	ns
CST English-Language Arts 10	48	367.56	37.60	373.36	4.32	9	393.56	45.74	362.63	10.84	683.90	0.80	ns
Math Group 2 (Algebra 1 / Geometry)													
CST Math 10	32	336.22	58.50	338.08	6.91	34	317.18	38.88	315.43	6.71	8424.90	5.52	0.05
CST English-Language Arts 10	33	381.24	28.82	379.42	3.77	34	368.56	40.78	370.33	3.72	1377.21	2.94	ns

Figure IX.15 Bar Graphs of 9th Grade (Baseline) Mean CST Mathematics and English-Language Arts Scores for Each Class by Intervention Status



Matched-Class Comparisons on Mathematics

Beginning with the results for the matched-class comparisons on Mathematics, only two groups were able to be constructed, as just described. One of these, which will be labeled Math Group 1, comprised students who took Geometry in 9^{th} grade and Algebra 2 in 10^{th} grade and whose CST Math 9^{th} grade mean class test score was above 325 points. The second group—Math Group 2—comprised students who took Algebra 1 in 9^{th} grade and Geometry in the 10^{th} grade and whose CST Math 9^{th} grade mean class test score was above 325 points. It was not possible to construct matched-class groups beneath the mean class Math test score of 325 points, as there were insufficient numbers of students in all except one class in the experimental group who took either of the Geometry/Algebra 9^{th}/10^{th} grade subject matter combinations.

Math Group 1

Moving to the results for Math Group 1, of the 14 classes from which students in this group were drawn, 4 classes (28.6%) were in the intervention school and the other 10 (71.4%) were in the control school. Altogether, there was a total of 129 students in this sub-sample, of whom 69 (53.5%) were in the experimental group and 60 (46.5%) were in the control group.

Clearly apparent in the results of the ANCOVA for Math Group 1 (see Table IX.15) is that the mean test score in 10^{th} grade CST Math for the students in both the experimental and control groups is virtually identical (359.71 versus 360.58, $p = 0.891$, not significant). What is interesting is that when the Mathematics score is held statistically constant in this way, the mean test score in 10^{th} grade CST ELA is significantly higher for the experimental group—by a margin of approximately 10 points—than it is for the control group (413.44 versus 402.96, $p < 0.05$). To the degree that we can regard this statistical control on mathematical ability as a way of approximately matching students on their intellectual ability, it would appear that the better ELA test performance of the students in the experimental group is unlikely to be explained by a difference in intellectual ability of students in the two groups. While this result is consistent with the expected effects of the TestEdge intervention, it will be necessary to show that this improvement in test performance is also associated with a reduction in test anxiety; this, indeed, is the case, as we will see shortly.

Math Group 2

For Math Group 2 (Table IX.15), of the 16 classes from which students in this group were drawn, 4 (25%) were in the intervention school and the remaining 12 (75%) were in the control school. In total, there were 67 students in this sub-sample, of whom 33 (49.3%) were in the experimental group and 34 (50.7%) were in the control group.

The results of the ANCOVA for Math Group 2 (Table IX.15) show that the mean test score in 10th grade CST Math is significantly higher for the experimental group—by a margin of approximately 23 points—than it is for the control group (338.08 versus 315.43, $p < 0.05$). It can also be seen that the approximately 9-point difference in mean test score in 10th grade CST ELA, favoring the students in the experimental over the control group, is likely the result of chance (379.42 versus 370.33, $p = 0.091$, not significant). However, because the ANCOVA did not also show a correspondingly greater decline in test anxiety in the experimental group when compared to the control group (TAI-Global, 2.27 versus 2.44, $p = 0.278$; not shown), we left further investigation of Math Group 2 for a later time.

Matched-Class Comparisons on English-Language Arts

Moving to the results for the matched-class comparisons on 9th grade CST ELA mean class score, only two (high-scoring and mid-scoring groupings) of the three matched-class groupings could be constructed; the third grouping (beneath a mean score of 320 points) had insufficient case counts of students in the control group for a valid comparison. The first grouping—ELA High Score—comprised classes with a mean test score above 375 points. Altogether, there were a total of 121 students in this sub-sample, of whom 69 (57.0%) were in the experimental group and 52 (43.0%) in the control group. The second grouping—ELA Mid Score—comprised classes with a mean ELA test score between 320 to 374 points; there were a total of 57 students in this sub-sample, of whom 48 (84.2%) were in the experimental group and 9 (15.8%) in the control group. While on both of these matched-class comparisons the experimental group's 10th grade ELA mean score is higher than that for the control group (413.23 versus 404.63, and 373.36 versus 362.63, respectively), the difference in neither of these results reaches statistical significance.

Math Group 1: Social and Emotional Correlates

Given the finding of a significant increase in CST ELA test performance in Math Group 1 of the experimental group, we wanted to identify the distinguishing sociodemographic characteristics of these students. Table IV.16 presents descriptive data for the students in Math Group 1, broken down by intervention status (percents by row).

Comparing the within-groups distributions (percents by column; not shown), in terms of Gender, the ratio of males to females was 47% to 53% in the experimental group, compared to 34% to 66% in the control group. While most (54%) students were White in the experimental group, the largest category in the control group was those in the Other Ethnicity category (49%). In both the experimental and control groups most students lived with both of their biological parents (83% and 68%, respectively). And while all 70 students in the experimental group were in an advanced academic class, about one in eight (13%) of those in the control group were in a regular class.

Table IX.16 Social Characteristics of Math Group 1 by Intervention Status

	Experimental		Control	
	N	%	N	%
Males	33	62%	20	38%
Females	37	49%	39	51%
White	37	60%	25	40%
Hispanic	22	81%	5	19%
Ethnicity Other	10	26%	29	74%
2 Bio Parents	58	59%	41	41%
Other Parents	12	39%	19	61%
Advanced Class students	70	57%	52	43%
Regular class students	0	0%	8	100%
White females	16	46%	19	54%
Hispanic females	13	87%	2	13%
Other females	7	28%	18	72%
White males	21	78%	6	22%
Hispanic males	9	75%	3	25%
Other males	3	21%	11	79%

The final step in the analysis of Math Group 1 was to investigate the relationship of the improvement in CST ELA test performance to test anxiety change and to changes on the SOS scales, both for Math Group 1 as a whole and also for subgroups of students with different backgrounds and family and class circumstances. Of the Gender, Ethnicity, Family Composition, and Class Academic Level characteristics examined in the analysis above, the categories of Female, White, Hispanic, Two Parent family, and Advanced Class were those for which a significant relationship with changes in test anxiety and/or the other SOS scales was found (see Table IX.17).

Starting with the results for the whole Math Group 1 sample, compared to the control group, there is a significant reduction in test anxiety in the experimental group (TAI-Global: 1.99 versus 2.22, $p < 0.05$) coupled with an increase in Positive Class Experience (3.08 versus 2.83, $p < 0.05$), which are associated with the higher mean CST ELA test score (413.44 versus 402.96, $p < 0.05$). Looking at the results for the other groupings of students within the Math Group 1 sample, a greater reduction of test anxiety in the experimental group is observed for Females (TAI-Global: 2.07 versus 2.35, $p < 0.05$), for students living with Two biological Parents (TAI-Global: 2.04 versus 2.32, $p < 0.01$), and for those in an Advanced Class (TAI-Global: 1.98 versus 2.23, $p < 0.05$).

An even stronger pattern favoring the experimental group over the control group across the characteristic groupings is evident for Positive Class Experience. With the exception of results for Advanced Classes, this is consistently higher in the experimental group for the other four categories—for Females (3.03 versus 2.70, $p < 0.05$); for White (3.15 versus 2.77, $p < 0.01$); for Hispanic (3.07 versus 2.08, $p < 0.01$); and for Two biological Parent family (3.14 versus 2.86, $p < 0.05$). Finally, for Hispanics there are two other significant correlates. One is the greater Parental Support (3.59 versus 2.76, $p < 0.05$) and the other is the more positive attitude about Life Preparedness (3.29 versus 2.58, $p < 0.01$) observed in the experimental group when compared to the control group.

Table IX.17 Math Group 1: ANCOVA of Post-Intervention Test Anxiety, Test Performance, and Selected Measures of Emotional Disposition for Whole Sample and Matched-Group Sub-Samples on Selected Characteristics by Intervention Status

	Experimental group					Control group							
Dependent	N	Mean	SD	Adjusted Mean	SEM	N	Mean	SD	Adjusted Mean	SEM	Mean Sq.	F	*p* <
All Students													
CST English - Language Arts 1	69	412.13	33.29	413.44	3.35	60	404.47	37.67	402.96	3.60	3505.71	4.53	0.05
CST Math 10	68	359.10	53.12	359.71	4.36	60	361.27	40.76	360.58	4.64	24.60	0.02	ns
Test Anxiety-Delta (Time 3)	60	-0.28	0.64	-0.30	0.07	54	-0.25	0.58	-0.23	0.07	0.14	0.52	ns
Test Anxiety-Global	70	1.94	0.74	1.99	0.06	59	2.27	0.82	2.22	0.07	1.61	6.13	0.05
Test Anxiety-Worry	70	1.87	0.79	1.88	0.07	59	2.09	0.82	2.07	0.07	1.18	3.68	ns
Test Anxiety-Emotional	67	2.03	0.81	2.09	0.08	58	2.45	0.99	2.37	0.09	2.43	5.73	0.05
Life Preparedness	70	3.11	0.71	3.09	0.07	60	2.99	0.63	3.02	0.07	0.14	0.46	ns
Parental Support	68	3.68	0.49	3.68	0.05	59	3.58	0.61	3.58	0.06	0.29	1.60	ns
Positive Class Experience	68	3.17	0.58	3.08	0.07	59	2.73	0.69	2.83	0.07	1.87	6.21	0.05
Extent of Friendship	70	3.01	0.66	3.06	0.05	59	3.25	0.56	3.20	0.05	0.63	4.16	0.05
Females													
Test Anxiety-Delta (Time 3)	31	-0.27	0.73	-0.28	0.10	37	-0.18	0.54	-0.18	0.09	0.16	0.53	ns
Test Anxiety-Global	37	2.01	0.78	2.07	0.08	38	2.41	0.72	2.35	0.08	1.49	6.19	0.05
Test Anxiety-Worry	37	1.90	0.86	1.92	0.10	38	2.20	0.76	2.19	0.10	1.34	3.93	ns
Test Anxiety-Emotional	36	2.16	0.86	2.23	0.11	37	2.61	0.93	2.53	0.11	1.62	3.71	ns
Life Preparedness	37	3.09	0.69	3.01	0.09	39	2.90	0.63	2.97	0.09	0.04	0.13	ns
Parental Support	36	3.70	0.44	3.72	0.06	38	3.67	0.52	3.66	0.06	0.07	0.56	ns
Positive Class Experience	36	3.18	0.59	3.03	0.09	38	2.56	0.68	2.70	0.09	1.79	6.32	0.05
Extent of Friendship	37	3.32	0.48	3.34	0.05	38	3.46	0.36	3.45	0.05	0.24	2.45	ns
White													
Test Anxiety-Delta (Time 3)	33	-0.31	0.51	-0.33	0.08	22	-0.23	0.56	-0.20	0.10	0.23	1.00	ns
Test Anxiety-Global	37	1.87	0.74	1.94	0.09	25	2.23	0.80	2.12	0.10	0.47	1.75	ns
Test Anxiety-Worry	37	1.82	0.72	1.85	0.09	25	1.98	0.83	1.93	0.10	0.09	0.32	ns
Test Anxiety-Emotional	36	1.91	0.85	1.99	0.11	24	2.45	1.04	2.33	0.14	1.64	3.63	ns
Life Preparedness	37	3.11	0.69	3.06	0.09	25	2.96	0.59	3.03	0.11	0.02	0.07	ns
Parental Support	35	3.76	0.37	3.77	0.05	25	3.76	0.42	3.76	0.06	0.00	0.01	ns
Positive Class Experience	35	3.24	0.46	3.15	0.09	25	2.65	0.73	2.77	0.10	1.93	7.95	0.01
Extent of Friendship	37	2.96	0.72	3.06	0.06	24	3.39	0.51	3.24	0.08	0.44	3.13	ns
Hispanic													
Test Anxiety-Delta (Time 3)	18	-0.46	0.53	-0.41	0.09	4	0.09	0.37	-0.13	0.20	0.24	1.65	ns
Test Anxiety-Global	22	1.99	0.73	1.97	0.09	5	2.25	0.82	2.37	0.19	0.66	3.67	ns
Test Anxiety-Worry	22	1.97	0.87	1.94	0.11	5	2.05	0.74	2.17	0.23	0.22	0.85	ns
Test Anxiety-Emotional	20	2.10	0.70	2.07	0.13	5	2.45	1.02	2.56	0.27	0.95	2.64	ns
Life Preparedness	22	3.30	0.53	3.29	0.09	5	2.53	0.30	2.58	0.18	2.09	12.75	0.01
Parental Support	22	3.56	0.68	3.59	0.14	5	2.90	1.15	2.76	0.30	2.71	6.36	0.05
Positive Class Experience	22	3.07	0.64	3.07	0.13	5	2.10	0.22	2.08	0.28	3.56	9.81	0.01
Extent of Friendship	22	3.08	0.46	3.06	0.09	5	3.00	0.76	3.09	0.20	0.00	0.02	ns
2 Parent families													
Test Anxiety-Delta (Time 3)	49	-0.29	0.68	-0.30	0.08	38	-0.24	0.55	-0.23	0.09	0.11	0.37	ns
Test Anxiety-Global	58	2.00	0.76	2.04	0.07	41	2.37	0.79	2.32	0.08	1.86	7.20	0.01
Test Anxiety-Worry	58	1.95	0.83	1.96	0.08	41	2.20	0.81	2.18	0.09	1.21	3.64	ns
Test Anxiety-Emotional	56	2.06	0.81	2.10	0.09	40	2.54	0.99	2.48	0.11	3.46	7.92	0.01
Life Preparedness	58	3.17	0.67	3.13	0.07	41	2.94	0.63	3.00	0.09	0.43	1.36	ns
Parental Support	56	3.75	0.38	3.76	0.05	40	3.68	0.58	3.66	0.06	0.24	1.81	ns
Positive Class Experience	56	3.21	0.52	3.14	0.07	40	2.76	0.64	2.86	0.09	1.64	5.93	0.05
Extent of Friendship	58	3.03	0.67	3.10	0.05	40	3.32	0.49	3.22	0.06	0.33	2.30	ns
Advancedd classes													
Test Anxiety-Delta (Time 3)	60	-0.28	0.64	-0.30	0.07	46	-0.21	0.61	-0.19	0.08	0.28	0.99	ns
Test Anxiety-Global	70	1.94	0.74	1.98	0.06	51	2.29	0.81	2.23	0.07	1.82	6.75	0.05
Test Anxiety-Worry	70	1.87	0.79	1.87	0.07	51	2.11	0.83	2.11	0.08	1.58	4.80	0.05
Test Anxiety-Emotional	67	2.03	0.81	2.09	0.08	50	2.46	0.95	2.38	0.09	2.45	5.78	0.05
Life Preparedness	70	3.11	0.71	3.09	0.06	52	3.08	0.61	3.10	0.08	0.01	0.02	ns
Parental Support	68	3.68	0.49	3.68	0.04	51	3.63	0.50	3.63	0.05	0.06	0.52	ns
Positive Class Experience	68	3.17	0.58	3.08	0.06	51	2.80	0.65	2.91	0.07	0.79	3.03	ns
Extent of Friendship	70	3.01	0.66	3.04	0.05	51	3.22	0.57	3.18	0.06	0.60	3.85	ns

ANCOVA

Summary of Results

The construction of matched groups using baseline CST scores led to the identification of two subgroups within the experimental group for which we found significant results:

- For Math Group 1 (9^{th} grade Geometry plus 10^{th} grade Algebra 2), a significant increase in CST ELA 10^{th} grade test performance was associated with a significant reduction in test anxiety.
- The higher test performance in Math Group 1 was also associated with an increase in Positive Class Experience for the whole sample as well as for females, Hispanics, and students from an intact family.
- Because the students were matched on mathematical ability, it is unlikely that the higher Math Group 1 test performance is due to a difference in intellectual ability between the experimental and control group students.
- For Math Group 2 (9^{th} grade Algebra 1 plus 10^{th} grade Geometry), a significant increase in CST Math 10^{th} grade test performance was found; it was not associated with a change in test anxiety.

Chapter X
The Question of Classroom and Teacher Effects

To this point, we have focused on students whom we have grouped according to their intervention status (experimental group versus control group) in relation to the sociodemographic variables, SOS scales, test anxiety, and test performance. In this section, we change our unit of analysis from the student to the class and investigate this question by grouping the student data by the 10^{th} grade homeroom class to which students were assigned. Overall, there were 36 10^{th} grade classes from both schools from which usable data were gathered; 21 classes were in the experimental group and 15 were in the control group.

Test Anxiety and Test Performance by Class

One question of importance, in relation to the efficacy of the TestEdge program, is the degree to which there is evidence that the intervention was effective in reaching all students, irrespective of the home class to which they had been assigned. Of particular interest is whether the TestEdge program has been developed to work more broadly in a variety of different classroom contexts, or seems more limited in scope and targets classrooms with better teachers and/or more academically motivated students. A measure of the reach and efficacy of the intervention is to examine its impact on test anxiety and test performance by class.

Beginning with the data for test anxiety, Figure X.1 presents two bar charts for each class—one chart for the experimental group (Figure X.1.a) and the other for the control group (Figure X.1.b)—showing pre- and post-intervention mean test anxiety, organized in descending order of pre to post-intervention change in test anxiety from left to right. Classes with a significant pre-post change in test anxiety are flagged with one or more asterisks (*) to indicate the level of statistical significance.

Beginning with the test anxiety results, 14 classes (66.7%) in the experimental group showed a significant reduction in the global measure of test anxiety (TAI-Global).[17] By comparison, only 3 classes (20.0%) showed a significant decline in test anxiety in the control group Figure X.1.b). This difference between the two groups in the incidence of classes in which a significant reduction in test anxiety occurred is statistically significant ($chi^2 = 5.87$, $p < 0.05$). Including all 16 classes in the experimental group flagged in Figure X.1.a for a significant reduction in test anxiety, either on the global test anxiety scale (TAI-Global) or on one of the sub-scales (TAI-Worry or TAI-Emotionality), it seems clear that the TestEdge intervention was effective in facilitating a reduction of test anxiety in more than three-quarters (76.2%) of the classes in the experimental group.

When the mean CST ELA test performance for each class is considered (shown in Figure X.2), what is especially noteworthy is that these reductions in test anxiety in the experimental group occur throughout the test performance spectrum, from low-scoring to high-scoring classes. For instance, significant reductions in mean test anxiety are shown in Class 20 and in Class 39; as is apparent in Figure X.2.a, these classes are the second highest and second lowest, respectively, in terms of mean CST ELA performance in the experimental group. These are quite remarkable results, given the less than optimal implementation of the Test Edge program by many teachers in the experimental school.

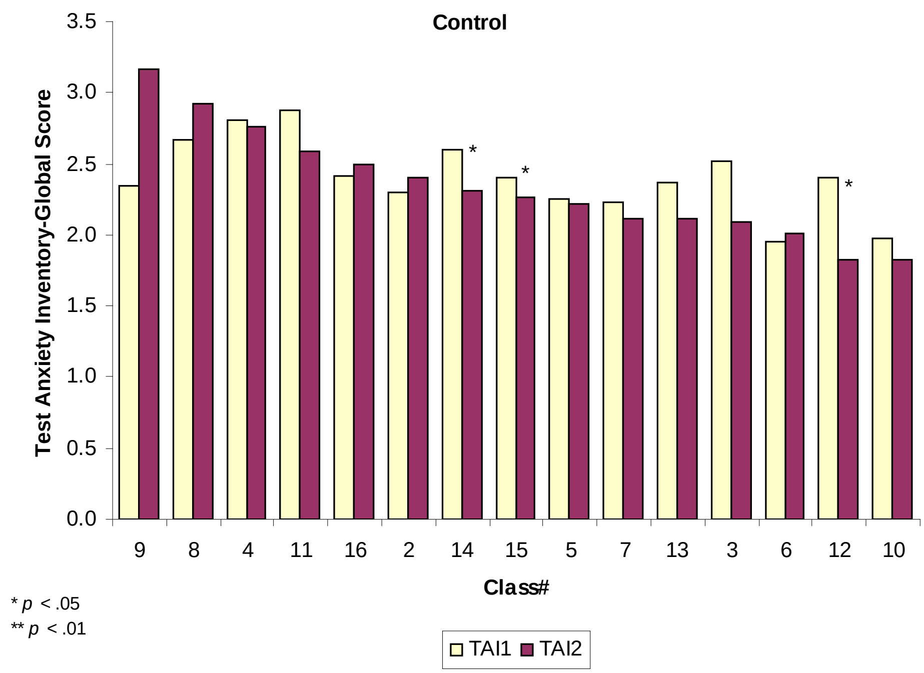
The results for test performance are presented in Figure X.2. Two bar charts are shown for each class—one chart for the experimental group (Figure X.2.a) and the other for the control group (Figure X.2.b)—showing the 9^{th} and 10^{th} grade CST ELA mean class scores, in descending order of 10^{th} grade score from left to right; classes with a significant change in mean score from the 9^{th} to 10^{th} grade tests are flagged with one or more asterisks (*), indicating the level of statistical significance.

[17] This result excludes the two classes in the experimental group that evidenced a significant reduction on the Test Anxiety-Worry sub-scale but not on the Test Anxiety-Global scale.

Figure X.1.a and X.1.b Bar Graphs for the Experimental and Control Groups Showing Pre- and Post-Intervention Mean Test Anxiety (by descending order of pre- to post-intervention change in test anxiety from left to right; Flags for significant change are shown).

Figure X.1.a

Figure X.1.b

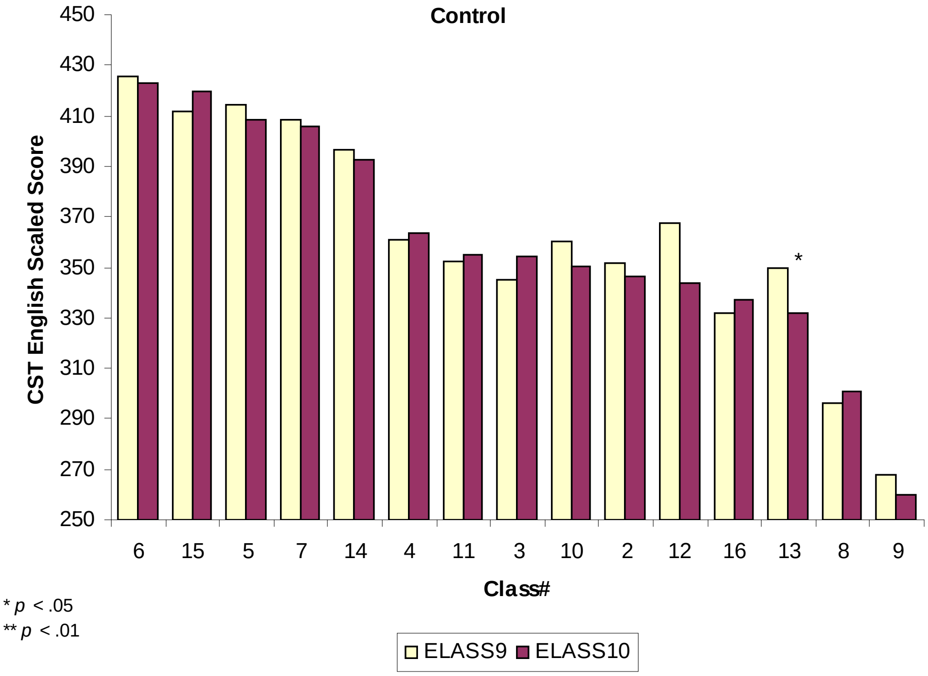


Examining the results for test performance, it is apparent that while 7 classes in the experimental group (33.3%) showed a significant change in mean class score, only two (9.5%) of these changes involved an increase in ELA score (Classes 20 and 17); a decrease was observed in the other 5 classes (23.8%) (Classes 33, 27, 26, 32, and 38), a result counter to the study's expectations. In the control group there was only one class (6.7%) with a significant test performance change, and this involved a decrease in ELA score (Class 13). These results show that while a significant reduction in test anxiety was achieved in a majority of the classes in the experimental group, such a reduction in test stress was not readily translated into an improvement in test performance.

Figure X.2 Bar Graphs for the Experimental and Control Groups Showing 9th and 10th Grade CST ELA Mean Class Scores (by 10th grade descending order; Flags for significant changes in test performance shown)

Figure X.2a

Figure X.2.b



* *p* < .05
** *p* < .01

Aggregating the class means by intervention status, we performed an ANCOVA to investigate pre-post intervention differences between the experimental and control groups on our measures of test performance, test anxiety, and the SOS scales. Given the markedly greater proportion of Hispanic students in the experimental group, two analyses were performed—one for all students (Table X.3), and the second without the Hispanic students (results not shown).

The results for all students show that while it is evident that the mean class CST ELA score is virtually the same in both groups (349.10 versus 348.96, p ns), classes in the experimental group have lower test anxiety (TAI-Global, 2.04 versus 2.34, $p < 0.001$), and also lower Emotional Discord (2.16 versus 2.27, $p < 0.05$) than classes in the control group. Conversely, the latter have a higher level of Interactional Difficulty than classes in the experimental group (2.03 versus 1.91, $p < 0.05$).

Removing the Hispanic students from the class means has little effect on test anxiety, which still remains significantly lower in the experimental group (TAI-Global, 1.97 versus 2.28, $p < 0.001$). However the differences on Emotional Discord and Interactional Difficulty are no longer apparent, which suggests that these are significant issues specific to the Hispanic student population.

Table X.3 ANCOVA of Test Anxiety, Test Performance, and SOS Scales by Intervention Status (Class Means)

	Experimental					Control							
Dependent	N	Mean	SD	Adj. Mean	SEM	N	Mean	SD	Adj. Mean	SEM	Mean Sq	F	*p* <
CST English - Language Arts 10	21	341.62	30.61	349.10	2.07	15	359.43	45.11	348.95	2.46	0.19	0.00	ns
Test Anxiety-Global	21	2.04	0.16	2.04	0.05	15	2.34	0.39	2.34	0.06	0.79	15.42	0.001
Test Anxiety-Worry	21	2.11	0.17	2.08	0.05	15	2.28	0.39	2.32	0.05	0.47	10.70	0.01
Test Anxiety-Emotional	21	1.97	0.21	1.99	0.06	15	2.39	0.39	2.36	0.07	1.12	16.40	0.001
Feelings about School	21	3.55	0.19	3.56	0.03	15	3.64	0.25	3.62	0.04	0.03	1.44	ns
Teacher Support	21	2.95	0.18	2.96	0.04	15	2.99	0.30	2.98	0.04	0.00	0.14	ns
Educational Plans	21	3.73	0.11	3.73	0.03	15	3.71	0.18	3.72	0.03	0.00	0.12	ns
Life Preparedness	21	3.05	0.16	3.02	0.03	15	2.92	0.16	2.97	0.04	0.01	0.66	ns
Parental Support	21	3.47	0.17	3.48	0.03	15	3.47	0.17	3.46	0.04	0.01	0.26	ns
Positive Class Experience	21	2.95	0.27	2.91	0.05	15	2.81	0.39	2.87	0.06	0.01	0.29	ns
Extent of Friendship	21	2.94	0.15	2.98	0.02	15	3.00	0.14	2.95	0.02	0.01	1.41	ns
Positive Affect	21	2.83	0.15	2.81	0.03	15	2.70	0.20	2.72	0.04	0.06	3.28	ns
Negative Affect	21	2.08	0.15	2.12	0.04	15	2.23	0.34	2.17	0.04	0.02	0.60	ns
Emotional Discord	21	2.13	0.18	2.16	0.03	15	2.32	0.19	2.27	0.04	0.09	4.78	0.05
Interactional Difficulty	21	1.89	0.15	1.91	0.04	15	2.06	0.25	2.03	0.04	0.12	4.57	0.05
Stress Management	21	2.42	0.14	2.39	0.04	15	2.44	0.22	2.48	0.04	0.06	2.34	ns

ANCOVA

Teacher Perceptions

As described in Chapter II, the socioemotional structure of the classroom can facilitate or impede student learning, growth and performance. Within the context of this social system, the teacher plays a pivotal role. The teacher sets the socioemotional tone; is responsible for creating and maintaining a safe, orderly, nurturing learning environment; and encourages, instructs, guides, and evaluates student academic performance. Thus, what the teacher perceives, believes, says, and does can have a significant impact on student perceptions, attitudes, expectations, and behavior.

Altogether, at baseline, there were 18 teachers included in the study from both schools, of which 9 were in the intervention school and 9 were in the control school. Our analysis began by comparing teachers from the two schools in relation to their perceptions and emotions as measured by some of the same items we used in the Student Opinion Survey.

The results for baseline measurement (Table X.4.a) show that there was no significant difference between teachers in the experimental and control groups on Test Anxiety-Global, Teacher Support, Life Preparedness, Positive Affect, Negative Affect, and Stress Management. Also, an ANCOVA run on pre-post changes found no evidence of a significant post-intervention difference between teachers in the two schools on these factors (see Table X.4.b). Because the teachers in the experimental group had completed the Resilient Educator Program in which they were fully instructed and trained in the TestEdge Program's emotional management tools, the lack of a post-intervention difference between the two groups of teachers, when taken at face value, is unexpected. However, given the reports and observations of a less than optimal implementation of the Test Edge program by many teachers in the experimental school (see Chapter XIII), the results are not surprising.

Table X.4.a ANOVA of Teacher Survey Scales at Baseline

	Experimental (N=9)		Control (N=9)			
	Mean	SD	Mean	SD	F	*p* <
Test Anxiety	2.14	0.73	2.06	0.50	0.08	ns
Teacher Support	2.78	0.33	3.11	0.55	2.40	ns
Life Preparedness	3.39	0.96	3.67	0.87	0.41	ns
Positive Affect	2.59	0.59	2.49	0.90	0.07	ns
Negative Affect	2.03	0.54	2.10	0.59	0.06	ns
Stress Management	2.63	0.33	2.76	0.58	0.33	ns

Single Factor ANOVA

Table X.4.b ANCOVA of Teacher Survey Scales—Post-Intervention

	Experimental (N=9)		Control (N=9)			
Dependent Variable	Adj. Mean	SEM	Adj Means	SEM	F	*p* <
Test Anxiety	1.89	0.22	2.03	0.22	0.20	ns
Teacher Support	3.08	0.14	2.88	0.14	0.97	ns
Life Preparedness	3.73	0.23	4.05	0.23	0.94	ns
Positive Affect	2.90	0.18	2.72	0.18	0.56	ns
Negative Affect	1.81	0.09	1.56	0.09	3.97	ns
Stress Management	2.80	0.18	2.72	0.18	0.11	ns

ANCOVA

When within-groups changes were analyzed (Table X.5), while there was some evidence of a mean reduction in test anxiety in the experimental group (before = 2.14, after = 1.92), statistical significance was not achieved, probably due to the small number of cases involved (N = 9). Despite this difficulty, however, a significant pre-post intervention increase in mean Positive Affect (from 2.59 to 2.92, $p < 0.05$) and a corresponding decrease in mean Negative Affect (from 2.03 to 1.79, $p < 0.05$) were found for teachers in the experimental group. In the control group, a significant decrease in Negative Affect was observed (from 2.10 to 1.75, $p < 0.01$).

Table X.5 Paired-Samples *t*-Test on Teacher Survey Scales for Teachers, Comparing Pre- and Post-Intervention Means in the Experimental Group and Control Group, Respectively

Experimental Teachers - Within-Groups Analysis

	Before (N=9)			After (N=9)					
	Mean	SD	SEM	Mean	SD	SEM	Difference	t	*p* <
Test Anxiety	2.14	0.73	0.24	1.92	0.63	0.21	-0.22	-1.24	ns
Teacher Support	2.78	0.33	0.11	2.96	0.51	0.17	0.19	1.17	ns
Life Preparedness	3.39	0.96	0.32	3.67	0.94	0.31	0.28	1.35	ns
Positive Affect	2.59	0.59	0.20	2.92	0.60	0.20	0.33	2.60	0.05
Negative Affect	2.03	0.54	0.18	1.79	0.34	0.11	-0.24	-2.29	0.05
Stress Management	2.63	0.33	0.11	2.78	0.48	0.16	0.15	1.20	ns

Paired Samples *t*-test

Control Teachers - Within-Groups Analysis

	Before (N=9)			After (N=9)					
	Mean	SD	SEM	Mean	SD	SEM	Difference	t	*p* <
Test Anxiety	2.06	0.50	0.17	2.00	0.88	0.29	-0.06	-0.21	ns
Teacher Support	3.11	0.55	0.18	3.00	0.50	0.17	-0.11	-1.00	ns
Life Preparedness	3.67	0.87	0.29	4.11	0.60	0.20	0.44	1.32	ns
Positive Affect	2.49	0.90	0.30	2.70	0.56	0.19	0.21	0.66	ns
Negative Affect	2.10	0.59	0.20	1.57	0.41	0.14	-0.52	-3.44	0.01
Stress Management	2.76	0.58	0.19	2.73	0.56	0.19	-0.02	-0.09	ns

Paired Samples *t*-test

Relationship Between Teacher and Student Perceptions

We turn next to the question of the relationship between teachers' perceptions and students' perceptions. Table X.6 presents three sets of Pearson correlation coefficients (pre- and post-intervention, and pre-post delta) showing the relationship between teachers' and students' (whole sample) perceptions on the six scales common to both the SOS and the Teacher Survey.

Looking over the correlations in the three tables, there is little strong evidence of a relationship between teachers' and students' perceptions, as there are only five correlations of statistical significance. While four correlations have a negative sign, there is one correlation with a positive sign. Interestingly enough, this is for the perception of pre-post change in Test Anxiety ($r = 0.60$, $p < 0.05$), and it suggests that student and teachers were in accord in their perceptions that student test anxiety had changed.

But for the four factors which have a negative sign—Pre Study: Stress Management ($r = -0.52$, $p < 0.05$); Post Study: Teacher Support ($r = -0.63$, $p < 0.05$) and Life Preparedness ($r = -0.55$, $p < 0.05$); and pre-post change: Negative Affect ($r = -0.65$, $p < 0.01$)—it would appear that teachers and students have opposing perceptions on these factors.

While these results are close to marginal significance, there is a pattern of consistency here which leads us to accept them as valid. The reader is cautioned, however, that it is possible—even likely, given the other findings above—that these aggregated results mask differences between the two schools which may have emerged had there been sufficient numbers of teachers for a robust breakdown of the correlations by intervention status.

Table X.6 Correlation Between Teacher and Student Perceptions on Scales Common to the Teacher Survey and Student Opinion Survey —Pre-Study, Post-Study, and Pre-Post Change

Pre-Study Teacher-Student SOS Correlations

		Teachers (n=15)					
		1	*2*	*3*	*4*	*5*	*6*
Students (N=749)	1. Teacher Support	-0.48					
	2. Life Preparedness	-0.50	-0.20				
	3. Test Anxiety	0.23	0.09	0.42			
	4. Positive Affect	-0.41	-0.12	-0.21	-0.22		
	5. Negative Affect	-0.05	0.02	0.01	-0.16	0.11	
	6. Stress Management	-0.52 *	-0.33	-0.24	-0.24	-0.25	-0.03

Post-Study Teacher-Student SOS Correlations

		Teachers (n=15)					
		1	*2*	*3*	*4*	*5*	*6*
Students (N=749)	1. Teacher Support	-0.63 *					
	2. Life Preparedness	-0.55 *	-0.44				
	3. Test Anxiety	-0.03	0.15	0.45			
	4. Positive Affect	-0.51	-0.40	-0.43	-0.09		
	5. Negative Affect	-0.04	0.28	0.43	-0.28	0.00	
	6. Stress Management	-0.08	-0.21	-0.33	-0.17	-0.10	-0.32

Pre/Post Change in Teacher-Student SOS Correlations

		Teachers (n=15)					
		1	*2*	*3*	*4*	*5*	*6*
Students (N=749)	1. Teacher Support	0.31					
	2. Life Preparedness	0.24	0.27				
	3. Test Anxiety	-0.06	-0.20	0.60 *			
	4. Positive Affect	-0.31	-0.38	-0.17	-0.09		
	5. Negative Affect	-0.40	0.01	0.02	-0.65 **	0.30	
	6. Stress Management	-0.13	-0.18	0.18	-0.23	0.10	-0.21

Summary of Findings

When students were aggregated by their home class our analysis of the data revealed the following results:

- A reduction in test anxiety was observed in more than three-quarters (76.2%) of the 21 classes in the experimental group. By comparison, only 3 classes (20.0%) showed a significant decline in test anxiety in the control group.

- The pattern of test anxiety reduction in the experimental group classes occurs throughout the test performance spectrum, from low-scoring to high-scoring classes.
- The results of the analysis of mean class test performance show that this reduction in test anxiety was not readily translated into an improvement in test performance in the intervention school.
- Aggregating the class means by intervention status, the results of an ANCOVA of pre-post intervention differences show that while class test performance is comparable in both schools, classes in the experimental group have lower test anxiety and lower Emotional Discord, whereas classes in the control group have a higher level of Interactional Difficulty.

In relation to the analysis of teacher perceptions, we found:

- No significant difference between teachers in the experimental and control groups on Test Anxiety-Global, Teacher Support, Life Preparedness, Positive Affect, Negative Affect, and Stress Management at baseline or at the time of post-intervention measurement.
- However, there was evidence of an increase in Positive Affect and a decrease in Negative Affect for teachers in the experimental group; a decrease in Negative Affect was observed for teachers in the control group.

From the correlation analysis of the relationship between teachers' perceptions and students' perceptions, we found that:

- While there is some evidence of concordance between teachers' and students' perceptions that student test anxiety had changed, on four factors—Stress Management, Teacher Support, Life Preparedness, Negative Affect—the evidence suggests that teachers and students have opposing perceptions.

Chapter XI
The Physiological Study

The primary goal of the TestEdge program is to teach students to more effectively manage their worry, fear, and anxiety when faced with a stressful situation, such as taking an important test, by making an intentional shift to a state of psychophysiological coherence. Assuming all other things are held constant, evidence of a post-intervention reduction in test anxiety in association with an increase in test performance, such as that we observed in Math Group 1, for example, would constitute compelling empirical documentation of a successful intervention effect. However, because the measure of test anxiety is subjective, based as it is upon self-reports by students in responding to the SOS questionnaire, corroborating data from an objective measurement such as HRV would provide an even stronger basis for such an inference.

To address this limitation of questionnaire data and provide an objective measure of post-intervention improvement in student stress management ability, we conducted an additional study, using electrophysiological instrumentation, on a randomly stratified sub-sample of students selected from both schools. This study was designed as a controlled laboratory experiment, with pre- and post-intervention measures of student ability to shift into the psychophysiological coherence state prior to performing a stressful computerized task.

Heart Rate Variability and Psychophysiological Coherence

The electrophysiological instrumentation throughout the experiment provided continuous measurement of changes in students' heart rate variability (HRV) and heart rhythm coherence. Heart rhythm coherence has been identified as the primary marker of the psychophysiological coherence state (McCraty et al., 2006). Moreover, various measures of HRV and coherence have previously been related to anxiety, emotional stability, cognitive function, behavioral inhibition, and task performance (Friedman & Thayer, 1998a, 1998b; McCraty et al., 2006; Pine et al., 1998; Porges, 1992a, 1992b; Porges et al., 1994).

As noted in Chapter II, heart rate variability is the natural fluctuation in beat-to-beat heart rate, which gives rise to what is commonly known as the heart rhythm. These beat-to-beat changes in heart rate are largely generated and amplified by the interaction between the heart and brain. This interaction is mediated by the flow of neural signals through the efferent (descending) and afferent (ascending) pathways of the sympathetic and parasympathetic branches of the autonomic nervous system (ANS). HRV therefore provides a measure of neurocardiac function reflecting heart–brain interactions and ANS dynamics that are involved in cognitive and emotional processes.

The mathematical transformation of the heart rhythm into power spectral density using Fourier transforms is used to discriminate and quantify sympathetic and parasympathetic activity as well as overall autonomic nervous system activity. Power spectral analysis reduces the patterns in the heart rhythm into its constituent frequency components and quantifies the relative power of these components. The HRV power spectrum is divided into three main ranges, and each range is associated with an underlying physiological mechanism that gives rise to the oscillations in that range.

The very low frequency (VLF) range (0.0033–0.04 Hz) is primarily an index of sympathetic activity, while power in the high frequency (HF) range (0.15–0.4 Hz), reflects more rapidly occurring changes in the beat-to-beat heart rate which are primarily due to modulation of the efferent parasympathetic activity associated with changes in respiration. The frequency range encompassing the 0.1 Hz region is called the low frequency (LF) range (0.04–0.15 Hz), and it reflects activity in the feedback loops between the heart and brain that control short-term blood pressure changes and other regulatory processes. The physiological factors contributing to activity in the LF range are complex, reflecting a mixture of sympathetic and parasympathetic efferent and afferent activity as well as vascular system resonance.

In our research on the physiological correlates of positive emotions (McCraty et al., 2006), we have found that when certain positive emotional states, such as appreciation, compassion, or care, are intentionally maintained, coherent heart rhythm patterns can emerge and be sustained, which reflects increased synchronization between multiple bodily systems.

Heart rhythm coherence is reflected in the HRV power spectrum as a large increase in power in the low frequency (LF) band (typically around 0.1 Hz) and a decrease in the power in the very low frequency (VLF) and high frequency (HF) bands. A coherent heart rhythm can therefore be defined as a relatively harmonic (sine-wave-like) signal with a very narrow, high-amplitude peak in the LF region of the HRV power spectrum and no major peaks in the VLF or HF regions. Coherence thus approximates the LF/(VLF + HF) ratio.[18]

Psychologically, the coherence mode promotes a calm, emotionally balanced, yet alert and responsive state. This state is conducive to enhanced cognitive and task performance. These enhancements include improved problem-solving and decision-making, as well as activities requiring perceptual acuity, attentional focus, coordination, and discrimination. Individuals generally experience a sense of enhanced well-being during coherence. This is due to the reduction in extraneous inner "noise" generated by the mental and emotional processing of daily stress and the positive emotion-driven shift to increased harmony in bodily processes. Many also report increased intuitive clarity and efficacy in addressing troublesome issues in life.

Experimental Protocol and Procedures

After students completed the SOS survey, an onsite analysis of the test anxiety scores was completed. Students from each class were divided into high and low test anxiety groups, and the research team randomly selected an equal number of potential student volunteers from the high and low test anxiety groups who were then offered the opportunity to participate in the additional study. In this way, the population for the physiological study was sampled across all of the classes in the school.

The sample population was randomly selected from the larger pool of students and an effort was made to stratify study participants into experimental and control groups by the following criteria: equal numbers of students in the high and low test anxiety categories; equal representation by gender; equal distribution among

[18] See McCraty et al. (2006) for a detailed description of the coherence mode and its quantification.

teachers and classrooms; and an equal representation of students in advanced and regular academic level classes. However, the effort to achieve equivalent experimental and control group samples was only partially successful. This was primarily the result of unanticipated conflicts in student class schedules with the prearranged time at which the physiological study was to be conducted.

After the pool of potential participants was identified, 136 subjects were randomly selected; we intentionally over-sampled well beyond our original sample target of 100 participants to provide a buffer for expected attrition over the intervention period. Lists of potential participants were created for each class, and research team members returned to the class to recruit students for the physiological study.

Students were offered two free movie passes to a local theater in exchange for their participation in the study: one pass for participation in the data collection that occurred prior to the TestEdge implementation, and a second for participation in the post-TestEdge program measurements. A signed parental Physiological Study Permission Form was required from each participant. Seventy-seven students from the intervention site and fifty-nine from the control site were recruited for participation in the physiological study.

Students were escorted from their respective classes by a research team member to a designated classroom where physiological measurements were obtained. Each student was asked if he or she was color-blind (this was necessary as the computer task they were to perform involved color recognition). It was also ensured that students were not chewing gum, as jaw movement would disrupt the physiological measurements. Each student was then assigned to a station equipped with a laptop computer to which was attached an earlobe sensor to measure his/her pulse. These pulse data were used to calculate heart rate variability (HRV) and heart rhythm coherence levels.

Each participant's pulse was recorded (Biopac MP 30) at a sample rate of 250 Hz during a resting baseline period of 5 minutes, which was followed by a stress preparation phase of 5 minutes. Students were told that the next phase involved measuring their speed and accuracy on a computer task and that they would be given an extra reward if they did well enough on the task, and so, for this 5-minute period (before the task), they needed to quietly prepare themselves. The participants were asked to use whatever techniques or strategies they normally used when preparing for an important test or challenge. This was followed by computer administration of four sections of the

Stroop Color-Word Conflict Test. Students were asked to do their best and told that their scores would be compared to those of students in another school. They were presented with instructions for each segment of the Stroop Test in written form on the computer screen as well as having the instructions read aloud by one of the researchers.

The post-intervention (Time 2) protocol for the experimental school site differed only in that the students were asked to practice one of the positive emotion refocusing tools they learned in the TestEdge curriculum during the stress preparation phase of the protocol. Students were reminded of the steps of the emotion refocusing technique by one of the researchers. For the Time 2 protocol at the control site, students again were asked to us their own methods to prepare themselves to take the Stroop Test.

The electrophysiological data gathered during the experiment were processed and analyzed at a later time at the Institute of HeartMath's Research Center laboratory, according to previously described protocols (McCraty et al., 2006; McCraty et al., 1999).

Physiological Study Sample

Table XI.1 presents descriptive data on the entire physiological study sample and compares the characteristics of students in the experimental school with those in the control school. Of the 136 students recruited for the physiological study, 77 (56.6%) were in the experimental school and 59 (43.4%) were in the control school. In terms of social characteristics, the experimental and control group samples were comparable in age (mean age in both was 15.3 years) and generally similar in family composition, with approximately 60% from an intact family living with both biological parents.

However, there were some differences in gender, ethnicity, and class academic level. Whereas there was an almost even division between males and females in the experimental group (53% and 47%, respectively), there was a greater proportion of females in the control group (60% versus 40%). Reflecting the ethnic characteristics of each school, 39% of students in the experimental group were Caucasian, 52% Hispanic, and 3% Asian; by contrast, in the control group 59% were Caucasian, 7% were Hispanic, and 24% were Asian. Finally, whereas 78% of the students in the experimental group were in a regular class, 63% of those in the control group were in an advanced class; this difference in academic class level emerged from the conflicts between the study's administration panels and many students' class schedules.

Table XI.1 Physiological Study Sample: Social Characteristics by Intervention Status

	Entire sample N = 136	Experimental N = 77	Control N = 59
Age, years (mean, sd)	15.3, 0.45	15.3, 0.44	15.3, 0.44
Gender			
Male	47%	53%	40%
Female	53%	47%	60%
Ethnicity			
Caucasian	48%	39%	59%
Hispanic	32%	52%	7%
Asian	12%	3%	24%
Other	4%	5%	3%
African American	1%	0%	2%
Pacific Islander	1%	0%	3%
American Indian	1%	1%	2%
Family Composition			
Both biological parents	61%	64%	59%
Single bio parent	15%	13%	17%
Mixed family, one bio parent	12%	14%	9%
Dual custody	9%	8%	10%
Relatives	1%	0%	2%
Other	2%	1%	3%
Class Academic Level			
Regular Class	60%	78%	37%
Advanced Class	40%	22%	63%

Baseline Equivalence

In relation to the baseline data on test anxiety, the SOS scales, and test performance in these two samples, there are some significant differences between the groups (Table XI.2). First, consistent with the difference in the primary study's whole sample, there is a large, significant difference of 46 points in baseline (9th grade mean CST ELA) test scores, with the students in the control group outperforming those in the experimental group (394.68 versus 348.33, $p < 0.001$). Also, the control group students are more positive in their Feelings about School (3.81 versus 3.44, $p < 0.05$) and report a greater Extent of Friendship (3.14 versus 2.87, $p < 0.05$) than the members of the experimental group. It is likely that these differences may be due to the higher proportion of students at the control school who were in an advanced academic class.

Table XI.2 Time 1 ANOVA of Test Anxiety, Test Performance, and SOS scales by Intervention Status

	Experimental Group (N=50)			Control Group (N=48)					
	Mean	SD	SEM	Mean	SD	SEM	Mean Sq.	F	*p* <
CST English - Language Arts 9	348.33	49.88	7.35	394.68	54.78	7.99	49952.99	18.18	0.001
Test Anxiety-Global	2.56	1.18	0.17	2.32	1.12	0.16	1.46	1.10	ns
Test Anxiety-Worry	2.60	1.20	0.17	2.30	1.12	0.16	2.18	1.62	ns
Test Anxiety-Emotional	2.52	1.20	0.17	2.33	1.17	0.17	0.87	0.61	ns
Feelings about School	3.44	0.66	0.09	3.81	0.74	0.11	3.27	6.60	0.05
Teacher Support	2.86	0.74	0.11	3.08	0.73	0.11	1.15	2.10	ns
Educational Plans	3.80	0.34	0.05	3.85	0.24	0.03	0.07	0.82	ns
Life Preparedness	3.10	0.65	0.09	3.01	0.68	0.10	0.18	0.41	ns
Parental Support	3.55	0.70	0.10	3.49	0.60	0.09	0.07	0.16	ns
Positive Class Experience	3.02	0.65	0.10	3.10	0.57	0.08	0.15	0.39	ns
Extent of Friendship	2.87	0.74	0.11	3.14	0.61	0.09	1.86	3.98	0.05
Positive Affect	2.87	0.67	0.10	2.74	0.65	0.09	0.41	0.92	ns
Negative Affect	2.23	0.71	0.10	2.19	0.67	0.10	0.03	0.06	ns
Emotional Discord	2.13	0.74	0.10	2.24	0.82	0.12	0.29	0.48	ns
Interactional Difficulty	1.91	0.61	0.09	1.90	0.54	0.08	0.00	0.00	ns
Stress Management	2.51	0.77	0.11	2.47	0.64	0.09	0.04	0.07	ns

Single Factor ANOVA

Statistical Analysis

Pre-Intervention (Time 1) Results

Table XI.3 shows the results of the physiological measures during the resting baseline period when students were sitting quietly waiting for the experiment to begin. With one important exception, Standard Deviation (SD) of RR (interbeat) Interval, there were no differences between the two groups of students on the physiological measures.

The standard deviation of interbeat intervals is a global measure of the overall amount of HRV. The lower SD of RR Interval in the experimental group (SD of RR intervals, 57.38 versus 66.18, $p < 0.05$) indicates that the overall amount of HRV in these students was lower than that of the control group. As was discussed in Chapter II, lower HRV is considered a psychophysiological marker of impaired emotional regulation and core regulatory functions in the development and maintenance of normal behavioral patterns. Although the majority of this research has focused on younger children and adults, one recent study with a population near the age studied here found that low HRV was associated with both externalizing and internalizing disorders (Pine et al., 1998). This suggests that the students in the experimental group were starting the study in a disadvantaged position.[19]

Table XI.3 Time 1 – ANOVA of Resting Baseline HRV Measures by Intervention Status

	Experimental Group (N=50)			Control Group (N=48)					
	Mean	SD	SEM	Mean	SD	SEM	Mean Sq.	F	*p* <
Heart Rate	79.38	9.70	1.37	77.22	10.43	1.50			
RR Interval	771.19	92.22	13.04	797.63	114.60	16.54	17121.62	1.59	ns
Standard Deviation of RR Intervals	57.38	20.07	2.84	66.18	22.42	3.24	1894.33	4.19	0.05
High Frequency Power	237.32	209.25	29.59	295.28	248.61	35.88			
Ln(High Frequency Power)	5.14	0.86	0.12	5.32	0.91	0.13	0.88	1.12	ns
Low Frequency Power	260.32	196.87	27.84	341.75	280.68	40.51			
Ln(Low Frequency Power)	5.31	0.72	0.10	5.55	0.76	0.11	1.41	2.58	ns
Total Power	586.21	449.31	63.54	745.15	488.77	70.55			
Ln(Total Power)	6.14	0.69	0.10	6.38	0.72	0.10	1.41	2.84	ns
Coherence Ratio	1112.32	7282.40	1029.89	88.23	199.51	28.80			
Ln(Coherence Ratio)	3.10	1.85	0.26	3.15	1.57	0.23	0.07	0.02	ns

Single Factor ANOVA

Table XI.4 Time 1 – ANCOVA of Stress Preparation Test HRV Measures by Intervention Status

	Experimental (N = 50)				Control (N = 50)						
Dependent	Mean	SD	Adj. Mean	SEM	Mean	SD	Adj. Mean	SEM	Mean Sq.	F	*p* <
Heart Rate	79.23	9.87	78.26	0.50	78.48	10.00	79.50	0.51			
RR Interval	773.21	92.29	784.84	5.02	784.82	106.70	772.71	5.12	3546.38	2.84	ns
Standard Deviation of RR Intervals	60.19	20.70	64.44	2.12	73.18	29.96	68.76	2.17	436.62	1.98	ns
High Frequency Power	220.97	198.14	244.40	24.68	300.02	305.48	275.61	25.20			
Ln(High Frequency Power)	5.09	0.82	5.16	0.07	5.26	0.97	5.18	0.07	0.00	0.02	ns
Low Frequency Power	290.50	198.43	316.53	38.40	483.65	395.73	456.54	39.20			
Ln(Low Frequency Power)	5.44	0.73	5.53	0.08	5.85	0.87	5.75	0.08	1.23	3.82	ns
Total Power	607.37	364.29	671.65	65.42	975.91	770.99	908.96	66.79			
Ln(Total Power)	6.23	0.63	6.33	0.06	6.57	0.85	6.46	0.07	0.43	2.15	ns
Coherence Ratio	556.14	3473.14	560.73	356.35	128.80	447.78	124.03	363.74			
Ln(Coherence Ratio)	3.22	1.75	3.23	0.23	3.16	1.63	3.16	0.24	0.13	0.05	ns

In the test preparation stage of the experiment, the students were instructed to "Do whatever you normally do when preparing for a stressful test." Table XI.4 presents the results of an ANCOVA of the students' HRV during this 5-minute period of the experiment in which the resting baseline HRV measures were used as the covariate to control for baseline differences. As expected, there were no differences between the

[19] In a correlation analysis (results not shown), it was found that low frequency power during the baseline recording was associated with performance on the CST Math 10 test. Students with lower LF power did not perform as well as the students with higher LF power (r = 0.246, *p* <0.05).

two groups during this phase of the experiment, since this was prior to the introduction of the TestEdge intervention in the experimental school.[20]

Post-Intervention (Time 2) Results

Following the TestEdge intervention, the experiment was repeated (Time 2) using exactly the same protocol as described for Time 1. As in the Time 1 protocol, the first phase in the experiment was a 5-minute resting period, during which the students were sitting quietly before the experiment actually began. The results of an ANCOVA conducted on pre-post changes in the resting baseline period are presented in Table XI.5.

There were a number of significant differences between the two groups. The first result of note is an overall increase in the experimental group's baseline heart rate variability, as indicated by the larger standard deviation of the interbeat intervals and increased Total Power (reflecting overall autonomic activity). The pre-intervention difference, favoring the control group, was reversed following the intervention, with the students in the experimental group showing now markedly greater HRV as compared to those in the control group (SD of RR Intervals 72.35 versus 55.47, $p < 0.001$; Total Power 1006.89 versus 501.72; Ln Total Power: 6.54 versus 6.00, $p < 0.001$). This indicates that an improvement in autonomic nervous system function has occurred in the experimental group.

Also of interest, as physiological indicators of the stress response, are the significant differences in mean heart rate, High Frequency Power, Low Frequency Power, and Coherence Ratio, all of which indicate improvements in the students in the experimental group over those in the control group. The mean heart rate was significantly lower in the experimental group (75.38 versus 79.62 BPM, $p < 0.01$; the indicator of parasympathetic activity (High Frequency Power) was higher in this group of students (Ln HF Power: 5.46 versus 4.93, $p < 0.001$); Low Frequency Power was also higher (Ln LF Power: 5.68 versus 5.15, $p < 0.01$), as was the ratio of heart rhythm coherence (Ln Coherence Ratio: 3.63 versus 2.79, $p < 0.05$). This last finding is also noteworthy as it indicates that the students have internalized the coherent state as a familiar physiological reference state. This objective physiological marker suggests that a substantial

[20]The correlation analysis also found a somewhat stronger association between the LF power during the stress preparation phase and academic performance on the CST Math 10 test. As in the baseline period, students with lower LF power did not perform as well as the students with higher LF power ($r = 0.397$, $p < 0.001$).

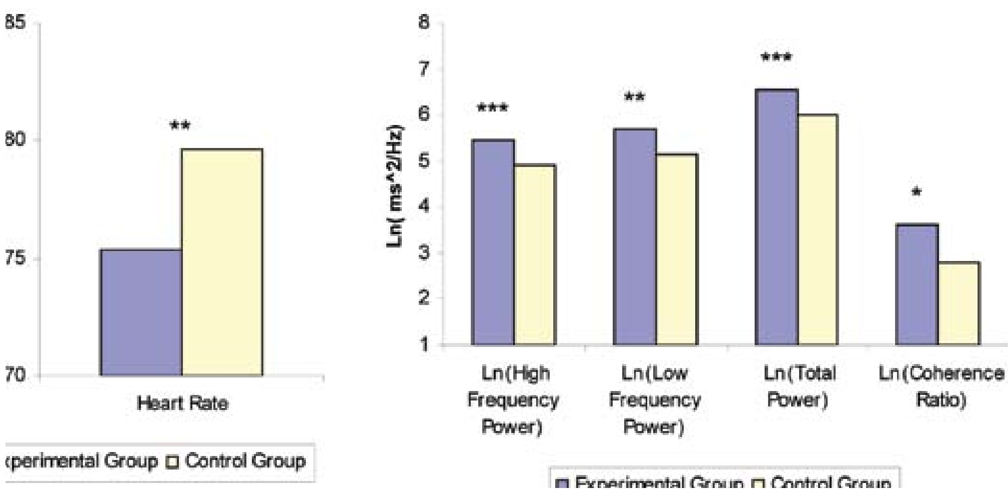
number of students had likely practiced the TestEdge coherence-building tools in their daily lives.

Table XI.5 Time 2 – ANCOVA of Pre-Post Resting Baseline Phase HRV Measures by Intervention Status

	Experimental Group					Control Group							
	N	Mean	SD	Adjusted Means	SEM	N	Mean	SD	Adjusted Means	SEM	Mean Sq	F	*p* <
Heart Rate	50	76.04	10.12	75.38	1.14	48	78.92	10.35	79.62	1.17			
RR Interval	50	809.03	106.71	817.27	11.59	48	777.89	102.51	769.31	11.83	55420.70	8.32	0.01
Standard Deviation of RR Intervals	50	68.80	30.91	72.35	2.92	48	59.17	21.78	55.47	2.98	6683.74	15.99	0.001
High Frequency Power	50	320.10	315.74	340.17	30.58	48	221.32	208.64	200.41	31.22			
Ln(High Frequency Power)	50	5.40	0.87	5.46	0.10	48	4.99	0.95	4.93	0.10	6.87	13.77	0.001
Low Frequency Power	50	499.01	838.14	538.63	80.99	48	262.91	197.34	221.63	82.69			
Ln(Low Frequency Power)	50	5.60	1.04	5.68	0.12	48	5.24	0.92	5.15	0.12	6.84	9.93	0.01
Total Power	50	925.07	1168.48	1006.89	104.00	48	586.94	387.88	501.72	106.17			
Ln(Total Power)	50	6.45	0.81	6.54	0.10	48	6.09	0.86	6.00	0.10	6.88	15.55	0.001
Coherence Ratio	50	144.37	286.48	145.68	30.70	48	48.98	98.11	47.61	31.34			
Ln(Coherence Ratio)	50	3.62	1.81	3.63	0.23	48	2.79	1.45	2.79	0.23	17.29	6.67	0.05

ANCOVA, Resting baseline HRV measure as covariate

Figure XI.5 Bar Graph Display of Adjusted Means in Table XI.5



For the stress preparation phase of the experiment, students in the control group were given same instruction as in the Time 1 experiment: “Do whatever you normally do when preparing for an important test or challenge.” Students in the experimental group were instructed to use one of the tools they had learned in the TestEdge program during this period to prepare themselves for the upcoming test. Table XI.6 and Figure

XI.6 show the results of an ANCOVA, using the Time 1 resting baseline HRV as the covariate.

On all measures of HRV, a significant difference was observed between the students in the two groups. Students in the experimental group had a lower mean heart rate (75.38 versus 79.62 BPM, $p < 0.01$) indicating they were less stressed, and they had greater high frequency power (Ln HF Power 5.46 versus 4.93, $p < 0.001$), indicating a higher level of parasympathetic activity, which is consistent with the lower heart rate. Their low frequency power was also much larger (809.23 versus 289.70; Ln LF Power, 6.16 versus 5.37, $p < 0.001$) which, when combined with the increased high frequency power, indicates that they were in a more relaxed, yet energized state associated with the psychophysiological coherence mode. This interpretation is confirmed by the significantly greater heart rhythm coherence ratio (Ln Coherence Ratio, 4.61 versus 2.79, $p < 0.05$), in the experimental site students.

Overall, the results from the physiological study present compelling objective evidence that the students in the experimental group had learned how to shift into the coherent state and better manage their emotions when preparing for a stressful task or situation, such as taking an important test.

Table XI.6 Time 2 – ANCOVA of Pre-Post Stress Preparation Phase HRV Measures by Intervention Status

	Experimental Group					Control Group							
	N	Mean	SD	Adjusted Means	SEM	N	Mean	SD	Adjusted Means	SEM	Mean Sq	F	*p* <
Heart Rate	50	76.84	10.33	76.21	1.24	48	78.96	10.80	79.62	1.27			
RR Interval	50	803.05	108.85	810.95	13.04	48	780.18	113.41	771.96	13.31	36622.26	4.34	0.05
Standard deviation of RR Intervals	50	78.63	32.57	82.38	3.15	48	63.71	24.05	59.80	3.22	11966.53	24.58	0.001
High Frequency Power	50	347.12	267.01	366.89	32.16	48	235.26	285.42	214.67	32.83			
Ln(High Frequency Power)	50	5.53	0.87	5.59	0.10	48	4.99	0.96	4.93	0.11	10.48	19.44	0.001
Low Frequency Power	50	772.95	1058.32	809.23	106.93	48	327.50	265.16	289.70	109.17			
Ln(Low Frequency Power)	50	6.10	1.02	6.17	0.13	48	5.44	0.92	5.37	0.13	15.06	19.65	0.001
Total Power	50	1248.80	1269.09	1328.81	122.62	48	711.58	541.81	628.23	125.18			
Ln(Total Power)	50	6.76	0.87	6.84	0.10	48	6.26	0.85	6.17	0.10	10.62	21.04	0.001
Coherence Ratio	50	364.21	594.18	367.62	60.40	48	30.24	36.51	26.69	61.65			
Ln(Coherence Ratio)	50	4.61	1.74	4.61	0.21	48	2.79	1.15	2.79	0.22	81.57	36.87	0.001

ANCOVA, Time 1 resting HRV measure as covariate

Figure XI.6 Bar Graph Display of Adjusted Means in Table XI.6

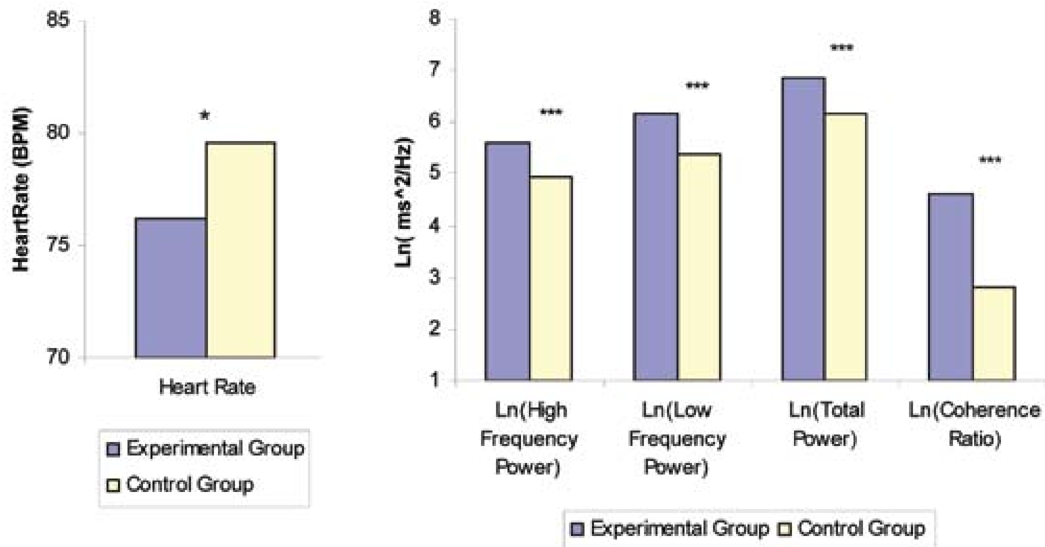


Figure XI.7 shows examples of the typical patterns observed in the HRV recordings collected during the stress preparation phase of the study at Time 1 (pre-intervention) and Time 2 (post-intervention). The four students whose HRV data are shown in the figure are typical examples drawn from the sub-sample of students matched on ninth grade CST Math scores, described above as Math Group 1.

In the figure, the students labeled 1 and 2 are from the intervention group (top) and the students labeled 3 and 4 (bottom) are from the control group. The left side of the figure labeled "Before" shows the students' heart rhythm patterns during the stress preparation phase of the experiment at Time 1, prior to the intervention. Next, under the label "After" are the heart rhythm patterns recorded during the stress preparation phase at Time 2, after students in the experimental group had learned the coherence-building tools taught in the TestEdge program. On the right side of the graph are their pre- and post-intervention test anxiety scores and CST–English-Language Arts test scores.

During the pre-intervention experiment at Time 1, the heart rhythm traces for all the students, especially students 1, 2, and 3, show the typical erratic and irregular pattern that we would expect to see during stress. Student 4's heart rhythm pattern has a higher degree of order, indicating that he or she may be utilizing an internal strategy such as rhythmic breathing to prepare for the upcoming Stroop Test. It can also be seen that the overall amplitude or range of the HRV waveform is smaller in the two students from the experimental school, which is in line with the pre-intervention HRV data from this group as a whole.

In the post-intervention HRV recordings, however, a clear shift can be seen in the heat rhythm patterns: the pattern associated with heart rhythm coherence is readily apparent in the students from the experimental site during the stress preparation period. It is also apparent that the amplitude of the HRV waveform has increased in these students, again mirroring the results found in the experimental group sample as a whole. By contrast, the Time 2 HRV recordings for the control students signify an ongoing incoherent psychophysiological state during the stress preparation phase.

Importantly, these examples also show that the students who learned to generate heart rhythm coherence had a corresponding reduction in test anxiety and an increase in test scores, while the control group students experienced an increase in test anxiety and reduced academic test performance, which again mirror the results for this subgroup of students as whole.

Figure XI.7 Typical Heart Rate Variability Patterns in Students Preparing for a Stressful Test

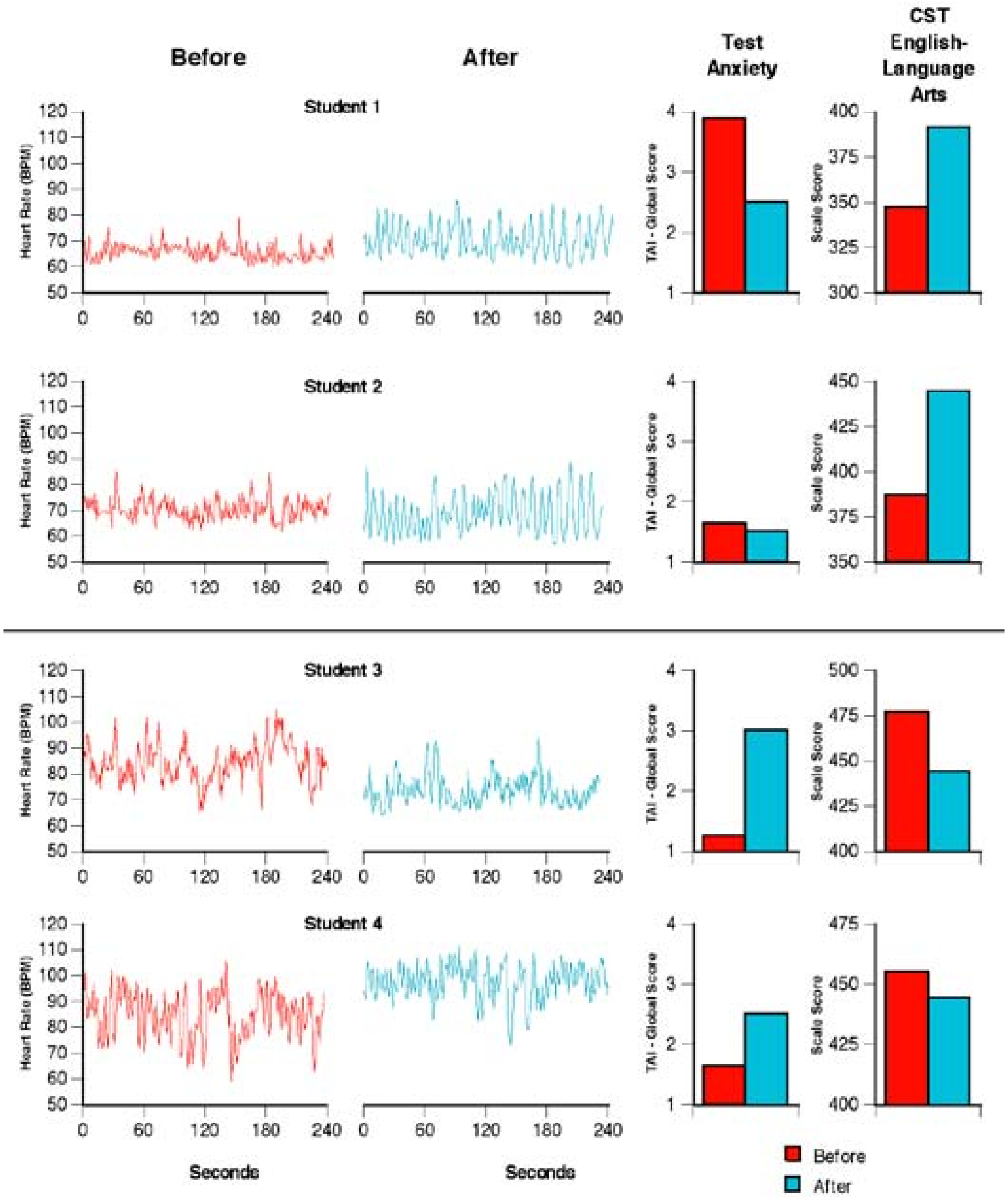


Figure XI.7. HRV recordings from the electrophysiological study, showing four students' heart rhythm patterns while they prepared for an important test, both before and after the TestEdge intervention. Pre- and post-intervention test anxiety level (TAI-Global Scale score) and the CST–English Language Arts test score for each student are also shown. For the two students in the intervention school, the recordings show a shift from an erratic, irregular heart rhythm pattern (left-hand side), before the intervention, to a sustained sine-wave-like pattern (increased heart rhythm coherence), indicative of the coherence state after the intervention. By contrast, both the pre and post HRV recordings for the students in the control school signify an ongoing incoherent psychophysiological state.

Test Anxiety, Test Performance, and Emotional Disposition

Table XI.8 presents the results of a correlation analysis of the relationship between the three test anxiety scales and student performance on the CAHSEE and CST at baseline, post-intervention, and pre-post change by intervention status. There are two interesting patterns in the results that differentiate the experimental and control group students.

The first is the pattern of correlations at baseline which shows a clear difference between the two groups. Whereas there are no significant negative correlations between test anxiety and test performance in the experimental group, there is a pattern of significant negative correlations between test anxiety (TAI-Global and TAI-Worry) and test performance (CAHSEE Math and 9th grade CST ELA) which range from -0.299 to -0.375 in the control group. This inverse relationship—the lower the test anxiety the higher the test performance—is consistent with the emotion-based theory of test anxiety outlined in Chapter II of this report. Why this relationship is not in evidence in the experimental group has not yet been identified in the analysis conducted to this point.

The second is the pattern of correlations for the pre-post changes in test anxiety and test performance, which is the obverse of that observed at baseline and points to a notable change both in and between the two groups. While there are now no significant correlations between text anxiety and test performance in the control group, a pattern of significant, moderate, negative correlations between the three test anxiety scales and 9th-10th grade CST ELA mean test score are evident in the experimental group, ranging from -0.344 to -0.389. This is clear evidence of a shift in the students in the experimental group. Since there is little evidence indicating an improvement in their mean pre-post test performance (see Table XI.9), this shift must result from the post-intervention reduction in test anxiety observed in the students in the experimental group.

Table XI.8 Correlations between Measures of Test Anxiety and Test Performance for Pre, Post, and Pre-Post Change by Intervention Status

	Experimental School			Control School		
	Test Anxiety Global	Test Anxiety Worry	Test Anxiety Emotional	Test Anxiety Global	Test Anxiety Worry	Test Anxiety Emotional
Pre Intervention Correlations						
CAHSEE MATH	-0.11	-0.163	-0.055	-.305*	-.352*	-0.249
CAHSEE English - Language Arts	-0.204	-0.244	-0.159	-0.224	-0.272	-0.17
CST English - Language Arts 9	-0.227	-0.26	-0.188	-.299*	-.375**	-0.214
Post Intervention Correlations						
CST English - Language Arts 10	-.285*	-.288*	-0.265	-0.268	-.350*	-0.177
Pre/Post Change Correlations						
CST English - Language Arts Change	-.389**	-.344*	-.350*	-0.019	-0.143	0.105

* $p < 0.05$ (2-tailed).

** $p < 0.01$ (2-tailed).

We turn, next, to the results of an ANCOVA (Table XI.9) controlling for baseline differences on the post intervention SOS scales and test performance for the students in the physiological study. Clearly evident are significant differences on all three measures of test anxiety, whereby mean test anxiety is lower in students in the experimental group than it is for those in the control group (TAI-Global: 1.94 versus 2.30, $p < 0.01$; TAI-Worry, 2.03 versus 2.29, $p < 0.05$; TAI-Emotional, 1.82 versus 2.29, $p < 0.001$). Also, Negative Affect is significantly lower in the experimental group as well (2.00 versus 2.35, $p < 0.01$). However, on the measures of test performance, 9th to 10th grade change in CST ELA, there is no significant difference between the two groups of students. This is not unexpected, as a higher ratio of the control group students were from advanced classes and had much higher tests scores at the beginning of the study (see Tables XI.1 and XI.2, respectively).

However, in an analysis of the relationships between test anxiety, coherence, and test performance in different subgroups of students in the physiological study, there do appear to be associations between improved physiological stability, and test performance when baseline academic ability is controlled, as we will see next.

Table XI.9 Time 2 – ANCOVA of Test Anxiety, Test Performance, and SOS Scales by Intervention Status

	Experimental Group					Control Group							
	N	Mean	SD	Adjusted Means	SEM	N	Mean	SD	Adjusted Means	SEM	Mean Sq	F	*p* <
CST English-Language Arts 10	46	341.70	46.30	360.85	4.65	47	391.55	57.49	372.80	4.60	2765.85	3.06	ns
Test Anxiety-Global	49	2.03	0.89	1.94	0.08	48	2.20	1.06	2.30	0.09	3.09	8.92	0.01
Test Anxiety-Worry	49	2.14	0.99	2.03	0.09	48	2.18	1.05	2.29	0.09	1.64	4.26	0.05
Test Anxiety-Emotional	48	1.88	0.85	1.82	0.10	48	2.23	1.13	2.29	0.10	5.38	11.09	0.001
Feelings about School	50	3.51	0.74	3.64	0.07	48	3.91	0.67	3.78	0.07	0.42	1.66	ns
Teacher Support	50	2.92	0.84	2.98	0.09	48	3.22	0.68	3.15	0.09	0.73	1.75	ns
Educational Plans	49	3.82	0.47	3.83	0.05	48	3.90	0.22	3.89	0.05	0.09	0.72	ns
Life Preparedness	50	3.06	0.68	3.04	0.08	48	3.06	0.66	3.09	0.08	0.07	0.24	ns
Parental Support	45	3.63	0.52	3.62	0.06	46	3.46	0.56	3.47	0.06	0.48	2.76	ns
Positive Class Experience	45	2.91	0.69	2.94	0.09	46	3.09	0.67	3.06	0.09	0.33	0.92	ns
Extent of Friendship	50	2.88	0.74	2.97	0.07	48	3.16	0.55	3.07	0.07	0.23	1.09	ns
Positive Affect	50	2.97	0.55	2.94	0.07	48	2.76	0.71	2.79	0.08	0.52	1.91	ns
Negative Affect	50	2.01	0.58	2.00	0.08	48	2.34	0.81	2.35	0.08	2.90	8.77	0.01
Emotional Discord	50	2.16	0.78	2.20	0.08	48	2.34	0.88	2.30	0.09	0.25	0.71	ns
Interactional Difficulty	50	1.82	0.60	1.82	0.07	47	1.94	0.56	1.95	0.08	0.37	1.37	ns
Stress Management	50	2.56	0.64	2.55	0.08	48	2.43	0.68	2.44	0.08	0.30	0.98	ns

ANCOVA

Matched-Pairs Analysis

In an additional analysis of the association between changes in the physiological parameters and test performance, we constructed a matched-pairs comparison by selecting students from the experimental group who conform to the expected relationship between test anxiety and test performance: we selected students who had a reduction in test anxiety and an increase in test performance. We then constructed a set of matched pairs by matching each student in this sub-sample with a student in the control group who had approximately the same baseline (9^{th} grade CST ELA) test score. We matched students to within a range of 5 test score points to each other, as a closer matching was not possible given the distribution of 9^{th} grade ELA scores in the two groups. This resulted in 11 matched pairs of students, including one matched-pair involving two students from the control group who had the same test score.

Table XI.10 shows the results of an ANCOVA comparing the matched-pairs on test anxiety, test performance, SOS scales, and heart rhythm coherence. As was found in the analysis of the entire sample, the experimental group students have significantly lower test anxiety (TAI-Global: 1.90 versus 2.41, $p < 0.05$) and significantly higher coherence (Ln-Coherence Ratio: 4.99 versus 2.78, $p < 0.01$). There is also a marginally significant difference of 9^{th}-10^{th} grade ELA test performance, in which the experimental group students outperformed their matched equivalents in the control group (365.54 versus 350.34, $p = 0.058$). In addition, the experimental group students had significantly lower Negative Affect (1.79 versus 2.53, $p < 0.05$) and Interactional Difficulty (1.55 versus 1.98, $p < 0.05$), and a significantly increased rating of Parental Support (3.68 versus 3.24, $p < 0.05$).

Table XI.10 Time 2 – ANCOVA of Test Anxiety, Test Performance, Three SOS Scales, and Stress Preparation Phase HRV Measures for Matched Pairs on CST ELA 9, by Intervention Status

	Experimental Group					Control Group							
	N	Mean	SD	Adjusted Means	SEM	N	Mean	SD	Adjusted Means	SEM	Mean Sq	F	*p* <
CST English-Language Arts 10	11	364.55	49.22	365.54	5.46	12	351.25	49.09	350.34	5.23	1323.87	4.04	=0.058
Test Anxiety-Global	11	1.75	0.93	1.90	0.18	12	2.54	1.16	2.41	0.17	1.50	4.39	0.05
Test Anxiety-Emotional	11	1.55	0.90	1.68	0.21	12	2.54	1.24	2.42	0.20	3.07	6.70	0.05
Parental Support	11	3.80	0.25	3.68	0.14	12	3.13	0.65	3.24	0.13	0.92	4.78	0.05
Negative Affect	11	1.68	0.48	1.79	0.19	12	2.63	0.72	2.53	0.18	2.55	7.28	0.05
Interactional Difficulty	11	1.47	0.24	1.55	0.12	11	2.05	0.42	1.98	0.12	0.62	5.38	0.05
Physiology Measures During Stress Prep Period													
Heart Rate	11	76.94	9.01	76.64	2.08	12	81.31	7.72	81.58	2.00	139.54	2.93	ns
RR Interval	11	797.22	92.24	801.50	21.34	12	749.68	71.21	745.76	20.43	17664.46	3.54	ns
Standard Deviation of RR Intervals	11	76.78	21.91	77.94	5.92	12	63.91	16.77	62.84	5.66	1228.52	3.30	ns
High Frequency Power	11	326.13	197.37	342.48	72.89	12	241.68	294.04	226.70	-	-	-	-
Ln(High Frequency Power)	11	5.54	0.85	5.55	0.25	12	5.05	0.91	5.04	0.24	1.47	2.08	ns
Low Frequency Power	11	644.92	603.69	637.98	136.91	12	315.04	198.20	321.41	-	-	-	-
Ln(Low Frequency Power)	11	6.11	0.90	6.13	0.26	12	5.53	0.78	5.52	0.25	2.07	2.87	ns
Total Power	11	1077.92	742.78	1112.86	182.46	12	660.08	433.29	628.06	-	-	-	-
Ln(Total Power)	11	6.75	0.77	6.77	0.21	12	6.31	0.64	6.29	0.20	1.29	2.65	ns
Coherence Ratio	11	381.56	437.57	382.64	91.87	12	34.87	51.95	33.88	-	-	-	-
Ln(Coherence Ratio)	11	4.95	1.83	4.99	0.46	12	2.83	1.14	2.78	0.44	27.51	12.13	0.01

ANCOVA

Discriminant Function Analysis

As a final step in our investigation of the physiological data, we conducted a pre- and post-intervention discriminant function analysis on the full sample of students in the physiology study. We were interested in investigating the degree to which post-intervention changes in test anxiety and psychophysiological coherence separated the experimental group students from those in the control group.

Discriminant function analysis is an ideal multivariate statistical procedure for investigating this type of question (Bradley, Young, Ebbs, & Martin, 1993). This is because the procedure aims to construct an additive linear model (the canonical discriminant function), composed of interval-level independent variables, that will maximize the separation (reduce the statistical association) between two or more nominal groups or categories which, together, are treated as the dependent variable. It also has the additional feature of providing a measure of the statistical model's predictive power by calculating the model's ability to classify cases into their correct *a priori* nominal grouping.

For our analysis, we used the nominal variable, intervention status, as the dependent variable and test anxiety, test performance, the SOS scales, and the measures of HRV as the independent variables. We conducted two sets of analyses: one on the data collected at the Time 1 measurement, prior to the intervention, and the second on the post-intervention (Time 2) data. And, to investigate the expected post-intervention changes in test anxiety and coherence, we ran separate analyses using the physiological data collected during the resting period and those collected during the stress preparation period. For all four analyses we used the subset of 98 students who had usable pre- and post-intervention data.

Pre-Intervention (Time 1) Results

The pre-intervention results for students during the resting period and also those for students during the stress preparation period were identical (results not shown). Of the twenty-four variables considered for entry into the statistical model in the stepwise procedure, only one variable, 9th grade CST ELA, had sufficient statistical power for inclusion (min. partial F to enter = 3.84; max. partial F to remove = 2.71). This is not unexpected, given the large difference between the two groups at baseline on the 9th grade CST ELA test.

However, the resulting canonical discriminant function produced only a small separation between the experimental and control groups (Wilks' Lambda = 0.848, Chi-square = 14.136, p <0.001), and only explained approximately 18% of the variance (Eigenvalue = 0.180, Canonical Correlation = 0.390). Even so, this model achieved a 67.7% correct classification rate in predicting student membership in their *a priori* groups; prior probability for group membership was 0.50.

Post-Intervention (Time 2) Results

From the analysis of the post-intervention data from the Stroop Test experiment, two discriminant function models were generated (results not shown), both of which were consistent with the expected effects of the TestEdge intervention articulated above. In a notable difference from the pre-intervention model, neither model contained the CST ELA test. Instead, the two models contained measures of test anxiety and HRV.

In more specific terms, the results for the resting period showed that two variables meet the criteria for entry in the stepwise procedure. Pre-post DS delta—a measure of the change in HRV parameters—was the first variable entered (Wilks' Lambda = 0.822), and pre-post test anxiety—measuring the change in test anxiety—was the second (Wilks' Lambda = 0.760). Together, they formed a canonical discriminant function which produced a modest separation between the experimental and control groups (Wilks' Lambda = 0.760, Chi-square = 21.962 p <0.001) and explained approximately 32% of the variance (Eigenvalue = 0.316, Canonical Correlation = 0.490). This model achieved a 70.1% prediction rate in correctly classifying students into their *a priori* groups; prior probability for group membership was 0.50.

The results of the discriminant function analysis for the stress preparation period were even more intriguing in that the changes from the resting baseline period to the stress preparation period were strongly consistent with the expected effects of the intervention. Three variables—two measures of HRV change and the measure of test anxiety change—met the criteria for entry in the stepwise procedure. The change in low frequency power was the first variable entered (Wilks' Lambda = 0.732), the change in test anxiety was second (Wilks' Lambda = 0.673), and the change in coherence ratio was the third (Wilks' Lambda = 0.641). Together, they formed a canonical discriminant function which produced a moderate degree of separation between the experimental and control groups (Wilks' Lambda = 0.641, Chi-square = 35.385, p <0.001) and explained approximately 56% of the variance (Eigenvalue = 0.561, Canonical Correlation = 0.599. This model achieved a 79.4% prediction rate in correctly classifying students into their *a priori* groups; prior probability for group membership was 0.50.

Overall, the results from the discriminant function analysis are highly consistent with the expected effects of the TestEdge intervention. By lowering test anxiety and increasing psychophysiological coherence in the students in the experimental group, it was expected that the effects of the intervention would shift these students away from students in the control group on these factors.

At Time 1, before the intervention, both during the resting baseline and the stress preparation periods prior to the Stroop Test administration, the only differentiator of the students in the two groups was performance on the 9th grade CST ELA test—a result consistent with the known difference in academic ability between the two schools.

However, by the time of post-intervention measurement, such test performance is no longer a common differentiator between the two groups of students. Instead, it has been replaced by changes in test anxiety and changes in the HRV measures of low frequency power and heart rhythm coherence, both of which are markers of the psychophysiological coherence state (McCraty et al., 2006; McCraty et al., 1999). While these factors are an effective means of differentiating students in the experimental and control groups during the resting period, they were even more powerful discriminator during the stress preparation period and were also able to correctly identify virtually 80% of the students who were exposed to the TestEdge program. In short, these results offer compelling evidence for the beneficial effects of the intervention on the students in the experimental group and corroborate the results presented in the prior chapters.

Summary of Findings

We conducted the physiological study to provide an objective measure of post-intervention improvement in student stress management ability. The study used electrophysiological measures and was conducted on a randomly stratified sub-sample of students selected from both schools. It was designed as a controlled laboratory experiment, with pre- and post-intervention measures of student ability to shift into the psychophysiological coherence state prior to performing a stressful computerized task. Results from the analysis revealed the following findings:

- Reflecting the sample bias, the control group sample had higher CST ELA test scores at baseline and were somewhat more positive about and more socially connected to their school
- There was lower overall heart rate variability at Time 1 (baseline) in the experimental group students, which, in relation to measuring a post-intervention improvement in HRV, placed these students in a disadvantaged position in relation to the control group; this was the only physiological difference detected prior to the TestEdge intervention.

The results from the post-intervention physiology experiment revealed a number of important findings:

- During the resting baseline period of the Stroop Test experiment there were a number of surprising improvements in HRV in the experimental group students—a marked increase in heart rate variability, and improvements in High

Frequency Power, Low Frequency Power, and the ratio of heart rhythm coherence—indicating that the students had likely practiced the coherence-building techniques and internalized psychophysiological coherence as a familiar reference state.

- During the stress preparation stage of the experiment there were significant differences between the two groups on *all* measures of HRV, indicating that students in the experimental group were less stressed, more relaxed, and more energized.
- These physiological differences were also associated with reduced test anxiety (on all three measures) and lower negative affect in the experimental group; however, there was no pre-post difference in test performance between the two groups.
- Results from an analysis of students in the experimental group who had both a reduction in test anxiety and an improvement in test performance, and who were matched on 9th grade CST ELA with students from the control group, show that they had lower test anxiety, lower Negative Affect and Interactional Difficulty, higher Parental Support, significantly improved HRV, and a marginally significant improvement in 9th-10th grade CST ELA test performance.

Finally, the results from discriminant function analysis, conducted to investigate the degree to which post-intervention changes on test anxiety and psychophysiological coherence separated the experimental group students from those in the control group, show:

- At Time 1 (before the intervention), both during the resting and the stress preparation periods of the Stroop Test, the only differentiator of the students in the two groups is student performance on the 9th grade CST ELA test—a result consistent with the known difference in academic ability between the two schools.
- By the time of post-intervention measurement, test performance was no longer a common differentiator between the two groups of students and had been replaced by changes in test anxiety and changes in heart rhythm coherence. Not only were these factors effective in discriminating between students in the experimental and control groups during the resting period, but they were an

even more powerful discriminator during the stress preparation period—the discriminant function constructed from these factors explained 56% of the variance and achieved a 79% prediction rate in correctly classifying students into their *a priori* groups.

In short, based on the objective data from the electrophysiological measurement, there is a strong pattern of consistent findings showing that students in the experimental group had learned how to shift into the coherent state and to better manage their emotions when preparing for a stressful task or situation, such as taking an important test. In addition, there is some evidence from a matched-pairs analysis that reduced test anxiety and increased coherence are directly associated with an improvement in test performance.

Chapter XII
The Face of Test Stress: A Study of Student Drawings

> *Sally (a teacher) shows me some of the artwork the students have created in their TestEdge workbooks … it takes my breath away. The booklet instructs the students to "Create a poem, story, song, cartoon, flow chart, or diagram," and suggests the students utilize one of several prompts related to the TestEdge curriculum, including "What is stress?" (Institute of HeartMath, 2002, p. 27). The students have poured their hearts and souls into these drawings, showing a level of vulnerability that I never observed in the classroom … .*
>
> —Laurie Schroeder (2006: 166)

How do students feel about high-stakes tests? A dramatic answer to this question is revealed in an unexpected way—through students' drawings.

In this section we report some preliminary but highly suggestive results from a partial analysis of data that are still being processed from the student drawings collected on the last page of the SOS questionnaire. The images students provided are a powerful visual testament as to how they feel about taking important tests. Moreover, these images also provide a window into how students see and feel about themselves and their classmates, teachers, and parents.

What follows are the results of analysis conducted to date on the limited data that have been coded from the student drawings database.

Research Background[21]

The primary purpose of this component of the study was to examine the relationship between student self-perception, as depicted in drawings of their experience of taking tests, and other quantitative measures of test anxiety, emotional disposition, physiological coherence, and test performance. The secondary purpose was to develop and test a new student drawing assessment designed by Laurie Schroeder (2004) for use by non-clinical researchers, teachers, and school counselors. In addition, this part of the study provided an alternative method for students to express their feelings about the high-stakes testing experience and for K–12 educators to further reflect on such experience. It was hoped that the data gathered from the student drawings assessment would not only cross-corroborate the results from the quantitative and qualitative analysis of data collected from other sources, but would also offer a new view of student feelings that would lead to a deeper understanding of the impact of high-stakes testing.

Most assessments used in educational research involve written language. However, as well documented in research on child development (see Bradley, 2001), because language is based on a system of abstract symbols and concepts, words are often an imperfect means of communication for children. This is even more the case if children have learning challenges or if assessments are not written in their home language, a common experience for non-English-proficient students (Schroeder et al, 2006:3). On the other hand, as summarized by Schroeder et al. (2006:2-3), art can often be a more direct means for students to express their perceptions, thoughts, and feelings.

> *As most educators know, art often provides an effective and enjoyable way for children and youth to "represent who they are, regardless of linguistic or cultural background" (Hasebe-Ludt, 1999, p.49). A natural mode of communication for children and often for youth, drawing can also, because of its non-verbal language, "be analyzed for structure, quality, and content" (Koppitz in Van Tillburg, 1987, p.6). Indeed, even unanalyzed drawings are an extraordinary source of information because the abstract and visual nature of the act of drawing brings "what has been below conscious awareness to a conscious level" (Lifford, Byron, Eckblad, Ziemian, 2000, p.56). This tends to discourage many students from using stereotyped or untrue images to manifest their feelings when they are asked to draw them.*

[21]This section draws heavily on Laurie Schroeder's work (Schroeder, 2006; Schroeder, Arguelles, & Bouman, 2006). Schroeder was a member of the TestEdge National Demonstration Study research team, and it was her idea to gather and analyze student drawings as a means of more directly accessing student perceptions and feelings about taking important tests.

When children draw, they form a "vibrant connection to the details of (their) worlds…we are invited to see and hear the marginalized tones, textures, and colors of their lives" (Mullen, 2002, p.12). In particular, students without strong verbal or written language skills can find an outlet in the kinesthetic process of creating a visual representation of what they are experiencing or have experienced.

There are many advantages to using students' drawings as an assessment tool in school-based research (Schroeder, 2006: 167-8). Art evokes deep-seated feelings in artists, observers, and researchers alike, crossing ethnic, cultural, gender, age, and language barriers. Drawings allow for the assessment of information and feelings that complement the data gained through more structured, cognitively oriented assessments. Through drawing, students without strong verbal or written skills can find an outlet for self-expression in the kinesthetic process of creating a visual representation of what they have experienced, currently are experiencing, or imagine they will experience. And while there are also a number of "pragmatic" advantages of using student drawings —a quick, non-intrusive, jargon-free, relatively open and unconstrained means of gathering students' expressions of their thoughts and feelings—these should be weighed against the enormous time and energy costs involved in coding and interpreting the drawings and analyzing the resultant data in relation to major themes and patterns.

Data Collection Procedure

The last page of the Student Opinion Survey (SOS) questionnaire was designed to collect student drawings. It contained a framed blank space in which students were instructed to draw a picture illustrating how they typically *feel* when taking an important test. Students were provided with a #2 pencil, and the prompt on the drawing page of the survey was read aloud by one of the researchers: "Please take a few minutes to draw or sketch a picture that shows how you feel when taking an important test (use the space provided below)."

A total of 830 pre-intervention drawings (535 from the intervention school and 295 from the control school) were collected in January, 2005, before classroom instruction of the TestEdge program had begun. Another 751 post-intervention drawings (468 from the intervention school and 283 from the control school) were collected in May, after students had completed the program. Altogether, a grand total of 1,581 drawings were collected on the SOS surveys given to the 10th grade students in the two high schools.

The results of two analyses, undertaken independently by two groups of researchers from the TENDS team, are presented here. The first, conducted by a group of CGU researchers, used an interpretative protocol developed by Schroeder (2004) to code and analyze themes and patterns identified in a random sample of 95 pre-intervention drawings. The second analysis was conducted by a group of Institute of HeartMath researchers using a different protocol they developed to code and analyze patterns of change in a sample of 163 pre-post pairs of drawings.

It should be noted that in the presentation of the results of the analysis of student drawings that follows, only a narrative discussion is offered; no images of the actual drawings themselves are presented in this report. This was done to ensure confidentiality and protect the identity of those individual students whose drawings we describe to illustrate the major themes and patterns identified in the analysis.

1. Analysis of Pre-Intervention Drawings

Interpretive Protocol

The drawings were analyzed utilizing the Schroeder Student Drawing Assessment Protocol (Schroeder, 2004), a holistic and interpretive instrument designed to systematically measure the degree to which certain specific content themes and patterns occur in student drawings. Designed for use by non-clinical researchers, teachers, and school counselors, this protocol consists of a set of standardized operational procedures for coding student drawings in response to a series of simple prompts. The measurement integrity of the Protocol was pre-tested in a pilot study conducted by Schroeder (2006).

The first section of the Student Drawing Assessment focuses on the observer's emotional response to each drawing. Observers were encouraged to use HeartMath tools to "attain physiological coherence, to improve their ability to remain in an emotionally neutral yet focused state" (Schroeder, 2004: 7), and also encouraged to be aware of their own biases and to set aside preconceived notions. Observers record their first overall impressions of the drawing, the emotions they evoke in them, overall perceived mood of the drawing, and other reactions or impressions about the drawing that they might have (Schroeder, 2004).

The second section of the assessment consists of a matrix to facilitate the process of coding each drawing. The design of this matrix was based on a coding scheme developed by Wheelock, Bebell, and Haney (2000a, 2000b). The matrix contains categories based on features shown in the student drawings in responding to a prompt eliciting their reflections about school and/or test-related experiences and feelings they have had or are currently having. Using the matrix, each drawing was coded individually and each applicable category was systematically recorded for every drawing. Examples of patterns and themes in the drawings include unambiguous and concrete content such as the figure of a student alone, the location of the student, or arrangements of desks in the classroom. They also include clearly discernible affective content such as facial expressions and body posture and other individual features such as the content of thought and speech bubbles and captions (Schroeder, Arguelles, & Bouman, 2006: 5).

The drawings were subjected to both quantitative and qualitative analysis, as described below.

Quantitative Analysis

Schroeder, Arguelles, and Bouman (2006) conducted a limited preliminary analysis of a sample of students' pre-intervention drawings, the results of which are reported in this section. They selected a random sample of 95 individual drawings which were then independently coded by three researchers using the Schroeder Student Drawing Assessment Protocol and rated for each applicable category. Inter-rater reliability was tested by computing the degree of intercorrelation among the independent ratings. Data from the affects construct were used to determine inter-rater reliability. The affects construct, developed to measure the valence of the emotion evoked in the coder viewing the student drawing, consisted of the following prompt: "The overall affect or mood of this drawing is: Positive, Neutral, or Negative." The three coding categories were signified by: 1 = negative, 2 = neutral, 3 = positive. The coding responses were tabulated for each rater and compared in a pair-wise manner. All correlations between the three pair-wise combinations were strongly positive and significant, and ranged from a low of 0.817 to a high of 0.919 (p <0.001, N = 95). Cronbach's alpha coefficient was also high, 0.945, indicating that the coding decisions made by the independent coders were highly consistent and statistically reliable.

Analysis of the affective mood or valence of the sample of drawings found that only 8% of the students feel positive when facing an important test, 23% feel neutral, and 69% feel negative (Schroeder et al., 2006: 6).

Qualitative/Interpretive Analysis[22]

The majority of the drawings convey intense, mostly negative emotions, as reflected in the quantitative analysis. The drawings, including those conveying intense negative affect, also reflect high levels of creativity and indicate that many students have highly developed drawing skills. In addition, humor is a component of many of the drawings, even those that depict very painful and angry themes. Schroeder et al. (2006) found a number of strong emotional themes in these drawings, as identified and illustrated in the narrative description of selected examples of student drawings in what follows.

Feelings of Isolation. In a large number of drawings students drew themselves alone, or, even when surrounded by other students or adults, conveyed a feeling of isolation due to stress and anxiety, as expressed verbally or by body language or facial expressions.

Feelings of Self-Diminishment. In a significant number of drawings, students drew themselves as disproportionately small. Many times, students drew themselves on the edge or on a corner of the drawing paper, a symbol of self-marginalization. In one moving example, a student drew a tiny box in the middle of the page, as if it were floating in space, with an even tinier image of him/herself inside the box. This image is a rudimentary stick figure, conveying an even more dramatic sense of anonymity and insignificance.

Feelings of Fear. In a majority of the drawings, students conveyed the many faces of the fear they experience when taking a high-stakes test. Some of the most common are: fear of disappointing parents, fear of not having enough time (as symbolized by the clock), fear of the unknown future, fear of expressing themselves, and fear of being judged and deemed unacceptable. For example, as depicted in another drawing, the fear of being judged is heightened by the student showing him/herself as being on fire. The sense of failure is heightened by a figure depicted pointing at the student and saying,

[22] This section is based on the work of Schroeder, Arguelles, and Bouman (2006) and is draws heavily from their paper.

"You failed. Ha, ha, ha, ha . . . ." Fear of ridicule—social shaming—is one of the strongest fears in adolescents. Some students added written comments to their drawings which expressed specific fears or which apologized for their perceived lack of artistic ability.

Feelings of Extreme Anxiety. A pervasive theme in the drawings is anxiety, specifically test anxiety. Students drew themselves biting their nails, sweating profusely, shaking, crying, and pulling their hair out. The depictions of body language in the drawings are richly descriptive of this anxiety: slumped shoulders, heads down on desks, rigid limbs, and staring at clocks. A striking illustration of extreme anxiety, fear, and a sense of being overwhelmed was contained in another example. In this drawing, not only the expression shown on the student's face, but his/her catalogue of negative physical feelings as well as the range of negative emotions: the flight/fight response ("I want to go home"), fear of disappointment ("Mom will be mad about my scores"), anxiety about the future ("Will I still be able to go to college?"), and feeling caught by the tyranny of the past ("I get bad scores every year") convey extreme mental and emotional stress resulting in a significant physiological overload: the student draws what appears to be an overworked, pounding heart (force lines radiate from it, accompanied by the word "bang!").

Feelings of Extreme Negativity or Emptiness. The drawings depict a preponderance of extremely negative facial expressions. Observers found the images truly haunting and unsettling. Students drew their eyes either laden with emotion or devoid of any emotion. Some drew their eyes as empty or as black holes. They drew their faces with expressions of intense anger, fear, confusion, and humiliation. One student who depicted himself as a jester drew eyes that, although almost closed, managed to convey intense anger and hostility. Only a handful of students drew faces that expressed confidence and happiness. Another drawing had what appeared to be a powerful allusion to Edvard Munch's famous painting, "The Scream," a graphic that has become an icon for modern existential angst and despair. In this drawing, a wide-eyed student holds both hands to his/her head while the elongated oval of their mouth forms what appears to be a scream of despair. One of their hands holds pencil with the pencil point pressed into his/her face in what looks like an act of self mutilation. On a desk, in front of the student lies a test with an "F" grade on it. In the background, a figure is drawn holding a fishing rod with the words "Future passing by" above.

Violent, Suicidal, and Other Feelings. Some of the drawings depict concrete violent, suicidal, and other destructive or self-destructive gestures and images. Sometimes such images and gestures are found in drawings with rudimentary, one-dimensional stick figures. In one such drawing, the student has turned the "hangman" game into a macabre joke, showing him/herself as being hanged in correctly responding "Y O U F A I L" to a fill-in-the-blank question. Another drawing depicts a much more violent image—a double suicide attempt: one hand holds a knife to the temple, the other pushes the barrel of a pistol right inside the mouth. A surprising number of drawings showed suicide ideation, self-mutilation, or other self-destructive feelings.

Evocative Metaphors

As a set, the drawings contain a number of rich, evocative metaphors, many of which Schroeder et al. (2006) categorized as having a "dark" theme. Some of the more prevalent and striking they identified were:

1. Hearts beating or exploding out of chests.
2. Students trapped as flies in spider webs.
3. Students ensnared in the vines of large plants.
4. Students with fire coming out of their mouths and backs, or being engulfed by flames.
5. Students boxed in or contained in bubbles.
6. Brains depicted as tornadoes or as gnarled masses of half-done math equations and history dates.
7. The Earth sitting squarely on the shoulders of a small stick figure.
8. Heads as empty containers.
9. Students falling or jumping off cliffs.

Of the various depictions of the latter theme, one image was particularly evocative. In this drawing, depicting what appears to be the futility of his/her enormous effort to pass the test, the student shows a steep, single-peaked mountain. On one side, a vertical arrow with the words "Climbing 2000 feet out of the Grand Canyon" points up from a second, lower arrow with the words "Represents all the study that I did." Just over the top on the other side, the student shows a small figure falling through space with the words "Me falling—Ahh!" Apparently, the fall does not represent the actual magnitude of failure on the exam but rather the extreme degree of the fear of failure the student feels. Pointing to the falling figure is an arrow with the words "Represents how freaked out I am so I feel like I'm going to fail."

In another powerful image, a student shows him/herself as a tortoise moving slowly down the middle of a highway lined with many clocks and timers, and posted with a sign: "SPEED LIMIT 300 MPH." All clocks and timers are ticking down to the zero hour, dramatically illustrating how completely the student will fail to complete the test on time.

There were also many drawings which combined a number of the themes described above. One example shows a student who is lying in bed with a clock and books and papers on the floor. Apparently s/he "Just fell asleep" and, despite having studied until 1:30 in the morning (the time shown on the clock), is now having a nightmare in which s/he is haunted by the realization that "I didn't do all my HW" (homework), is feeling overwhelmed ("too many bubbles," written in large letters) by the specter of three tests the next day ("When it rains it pours. Tests: math, science, english [*sic*]"), and anxious about his/her physical needs ("not enough sleep" and "I need to eat breakfast").

A second example is also crowded with images of despair. The student presents herself as completely frazzled ("I'm <u>tired</u>!"), with large, blankly staring, bloodshot eyes and with tears (or possibly sweat) streaming down her face. The bubbles all around her represent her dark thoughts—"I hate bio"; "I'm going to fail"; "I'm not going to get into college." She depicts the ultimate consequence of not passing the test—death—as represented by her gravestone which reads, "The Girl who couldn't pass the test." Even further indication of her tenebrous mood is the fact that it is nighttime in the cemetery—she draws the moon and stars around her gravestone.

Depictions of Adults

Significantly, there were very few depictions of adults in the drawings. Of the handful of images, the vast majority were negative, depicting adults as oblivious or insensitive to the students' feelings, or even as threatening to students. Two examples are illustrative.

One image shows a student about to be crushed by a giant pencil with a broken tip, behind which stands a teacher who says calmly, "It's just a test." The contrast between the two sides of the drawing reveals the difference between the student's and the teacher's perspectives. On the one side, the 45-degree angle of the student leaning away and shouting "No" to the threatening pencil appears to illustrate fear and tension, while on the other side the upright figure of the teacher holding a "test" seems to symbolize calmness and order. This difference is also illustrated in the contrasting look on the faces of the two figures, with the student's face revealing terror and the teacher's face reflecting confidence and composure. An even further contrasting detail is the difference in the two figures' eyes, with the student's eyes wide open in horror and the teacher wearing dark glasses, suggesting that she is blind to the student's feelings about tests. A final telling detail is the broken pencil tip, suggesting that even if the student knew the answers, she couldn't write them down.

Another drawing depicts a more ominous relationship between teacher and student. In this stark drawing the student sits at a desk taking a test, almost like an automaton, while the teacher stands behind holding a gun to his/her head. Behind the student, in the background, as if to reinforce the futility and hopelessness of the student's predicament, is the key to the test, which floats above the teacher's desk like a tantalizing but inaccessible talisman that the student can't see.

One possible explanation for the paucity of adult representation in these drawings is that these students find themselves in a developmental stage characterized by the process of psychosocial differentiation from adult role models. However, one can also entertain the possibility that the intense negative experience of test-taking may lead to an increase in the marginalization of youth from the adult social world. This may particularly include minority youth who are still part of an intergenerationally integrative culture (Schroeder et al., 2006: 8).

Written Embellishment

Also significant was the fact that many of the drawings were accompanied by written text, as if the images alone were not sufficient to convey the thoughts and emotions student's associated with test-taking. This expression took many forms: thought and conversation bubbles, signs on walls, captions, and titles. Some of the written comments were humorous, while many were heart-breaking: "Help me!"; "I am tired";

"Why do they stress us with all this?"; "My parents are going to kill me!"; "I am not going to get into college!"; "Suicide!"

Positive Feelings

While the majority of student drawings conveyed negative emotions and sentiments, a few depicted positive feelings about tests. Most of these tended to be simpler illustrations, without graphic adornment or a multiplicity of written messages. One example is a simple illustration of a student solving a test problem. A stick figure sits at a table with a test in what appears to be a composed thoughtful pose, scratching his/her head with one hand in a gesture that says "I'm thinking" as s/he strives to recall or discover the answer. Another example is a face, drawn in simple, direct lines, with a focused "matter-of-fact" expression (eyes wide open and pupils looking directly in the eyes of the reader) and with the word "CALM!" written underneath in capital letters, with an exclamation mark, and underlined to emphasize his (he has crew-cut hair) response to tests. A third example is depicts a smiling-faced girl (she is wearing a dress and has long curly hair) standing in a grassy field near a tree, while birds and soft billowy clouds float overhead; a page with the word "TEST!" is drawn in the foreground. Although she acknowledges in a thought balloon that she is "trying too hard," she still seems to have a positive, confident stance toward the test: "I'm kind of confused but I'm not scared or nervous," is written in the second thought balloon. She has cleverly used the same symbol ("?") for the test questions as she has for her eyes.

2. Analysis of Pre–Post-Intervention Change

Quantitative Analysis

Following up on Schroeder et al.'s (2006) analysis of a random sample of 95 pre-intervention drawings (reported above), Sam Bouman, a CGU graduate student and member of the research team, conducted a pre-post analysis of the drawings from 661 students. These drawings were evaluated in relation to Schroeder's affect construct and independently coded by three researchers following the exact procedures described above. (The affects construct was developed to measure the valence of the emotion evoked in the coder viewing the student drawing and consisted of the following prompt: "The overall affect or mood of this drawing is: Positive, Neutral, or Negative." The three coding categories were signified by: 1 = negative, 2 = neutral, 3 = positive.)

An analysis of the difference in means, using a *t*-test, was performed to see if there was a significant difference between the pre-intervention mean affect construct rating between the intervention and control groups (mean = 1.51 and 1.58, respectively). The mean difference (0.07) was not significant (*t*-test, -1.208; $p > 0.05$) indicating that, as expected, the affective valence of the student drawings in the two schools was similar before the TestEdge program was implemented.

Comparing the two schools on the affect construct ratings after the TestEdge intervention revealed that while only a small mean increase of 0.07—slightly more positive affect—had occurred in the intervention group, a larger decrease of -0.23—somewhat more negative affect—was observed in the control group. A *t*-test of difference in means between the schools (0.23) found that this post-intervention change was larger than expected by chance (*t*-test, 3.767; $p < 0.01$). To test the observed difference between the intervention and the control groups, while controlling for pre-intervention differences, an ANCOVA was performed. The ANCOVA results found that the differences were significant ($F = 16.839$, $p < 0.01$). Overall, these pre–post-intervention results suggest that while the raters observed a slight increase in positive affect in the drawings from the intervention school, there was a greater increase in negative affect in the drawings from the control school (Bouman, 2006).

These promising preliminary results will be further investigated in a subsequent analysis, expanded to encompass a full coding of the database of the 1,581 student drawings within a pre–post-intervention framework.

Qualitative Analysis

A group of HeartMath researchers conducted a preliminary qualitative analysis of changes observed in the drawings on a matched-group sample of 163 pre-post drawing pairs. The goal was to evaluate the potential impact of the TestEdge intervention on student self-perception by comparing the pair of pre- and post-intervention drawings for each student in two subgroups in the experimental school with comparable students in the control school. The two subgroups we used were those for which we had previously found a significant reduction in text anxiety in association with a significant improvement in test performance—White Females in a Regular Class and Math Group 1 (see Chapter 9 above).

The images in each student's pre-post pair of drawings were first examined for patterns of change along three dimensions—positive/negative changes in feelings and cognitions; positive/negative changes in depictions of self; and positive/negative changes in perceptions of self-control and success—and then classified into one of three general categories of change: those that appeared positive; those that appeared negative; and those that appeared to suggest no change.

Although this must be confirmed by a formal analysis, our initial results suggest that while approximately one-quarter of the students in each experimental school subgroup showed no change in their pre-post drawings, somewhat more than half of the students in each subgroup depicted a positive pre-post change. The remaining one-quarter evidenced a negative change between their pre and post drawings. The degree of positive change in student drawings from the intervention school appears be about twice as large as that observed in the pre-post drawings of the two control subgroups. What follows are a number of narrative descriptions of typical pairs of pre- and post-drawings showing positive change, selected from the intervention and control school sub-samples.[23]

Pre-Post Examples from the Intervention School

Positive Change in Self-Concept

Three examples of pre-post pairs of drawings have been selected to illustrate a positive change in a student's concept of self observed in the pre-post comparison of the drawings. The first pre-post pair example shows the movement from the pre-intervention image of a stark, square-headed figure with no neck or body and stick-like limbs, all alone—except for the sun—in an empty space, to the post-intervention image of a more realistic drawing of a fully-clothed person, with hair, ears, and a broad smiling mouth standing with a squirrel next to a tree under the sun. The second example depicts the change from a fearful, impotent student to one who appears confident and is in control. In the pre-intervention drawing, the student sees himself shrunk into inescapable insignificance in a tiny box, with the word "Me" pointing to it, under a huge wall clock. Above the clock is drawn a bottle of "Glue," from which a few drops fall as if in a fantasy to slow or stop the inexorable movement of time. By contrast, in the

[23]We plan to present all three categories of change in a subsequent publication upon the completion of our formal analysis of the student drawings dataset.

post-intervention drawing the student depicts the test as a source of nourishment: it is shown lying on a plate—with knife, fork, and spoon at the ready—as a meal he will consume.

The third example pair of drawings shows a change from dying in response the test to actually surviving it. In the pre-intervention drawing, set among tombstones in a graveyard with a leafless tree and an owl, the student depicts the test as a threatening weapon (with hands holding a club and a long knife) and herself as a small stick-figure who, in her own words will inevitably be "murred [murdered?] by the test!" On her tombstone , she carefully inscribes her name, her date of birth and her death date ("die on test day"). However, by the post-intervention drawing this student now seems more realistic both about her feelings and about the testing situation. Even though she sees herself (placed in the upper-right corner) as "hiding in the shadows," she now draws herself as a fully-clothed young woman, with long hair and bangs, and sad-but-pretty facial features with tears falling on her left cheek. And although she is still pessimistic about the test, which is shown lying on a table with an empty chair in the foreground, she seems much more in touch with the complex array of her feelings: "not good, sick," "intense," "sad," "forgetful," "scared," and "stress." Moreover, she even acknowledges that she is not alone, and that her teacher, although shown as distant (drawn facing her from the opposite corner), is encouraging her to "Come, sit, and good [do] your best." That's quite a change in four months!

Positive Change in Feelings and Sentiments

We have selected three example pairs of drawings to illustrate a positive change in a student's feelings and sentiments between the pre- and post-intervention period. The first pair of drawings shows a dramatic change from a dark, oppressively negative emotional state to one that is light, open, and positive. In the pre-intervention drawing, the student uses heavy, highly energized, bold strokes to depict a black sky crushing down on a delicately drawn leafless tree (the student?), which is surrounded by dark, lifeless, block-like objects all entangled on the ground by a wire-like coil. The words "<u>NERVOUS</u>" (capitalized and underlined) and "TOO MUCH PRESSURE" hang in the air near the tree. In the post-intervention drawing, light, flowing, whimsical strokes, depicting two large flowers with round petals under a bright sun, convey a strong positive feeling of optimism and life.

The second example illustrates a dramatic pre-post change in emotions, thoughts, and behavior. In both drawings the student shows herself sitting taking a test. In the pre-intervention drawing, the student portrays herself as tense and uptight (stiff-backed with lank hair), grimacing with her eyes closed in an agitated (tapping her foot), confused state. The thought balloons around her head indicate that she is more focused on her discomfort ("I'm uncomfortable"), the distractions around her ("Its [*sic*] too noisy"), and on negative mental ideation ("I don't understand"; "I forgot"; "I'm bored") than on performing well on the test. After the intervention, she shows herself taking the test with a positive, optimistic attitude, signified by her open, smiling face, and curled-up hair. She is no longer distracted by the test-taking circumstances. Instead, she is calmly focused on the task at hand (her eyes are open) and, while still acknowledging that "I'm bored," she is now finding that "this is easy." She also seems to have become more emotionally aware of her reactions to test-taking in that she observes, "I get different thoughts & emotions 4 [*sic*] every test."

The third pair of drawings exemplifies a positive change in emotions and sentiments, depicting the movement from a highly agitated emotional state to one of calm and serenity. In the pre-intervention drawing, the student is on her knees in what appears to be a state of panic, face lifted upward and praying/appealing to an external agent with an obsessive mantra: "Must Remember … Must … Must …" A thought bubble encloses a chaotic jumble of lines, suggesting that her mind is in total confusion. The multiple, heavier, squiggly pencil strokes of the drawing reinforce the sense of a highly energized, nervous emotional state. By contrast, the softer, delicate lines in the post-intervention drawing reinforce the sense of calm and serenity conveyed by the image. She depicts herself sitting at a desk about to take the test. She appears internally focused—her eyes are closed and hands are gently clasped together as in a prayerful state—with a serious expression on her face as she prepares to take the test that lies before her.

Positive Change in Perceived Self-Control and Success

The final two pairs of drawings we describe illustrate the third pattern of pre-post change observed in the drawings from the intervention school—a positive change in perceived self-control and visualization of success.

The first pair depicts a striking pre–post-intervention change from a single global image depicting a strong overwhelmingly negative, stressed-out state and the anticipation of failure to the student visualizing a sequence of ten specific images of herself taking the test and succeeding. In the pre-intervention drawing the student depicts herself as at the mercy of external forces—as having no control over her performance on the test. She holds her test and a pencil uselessly in each hand, while a thunderous storm is about to strike her with a lightning bolt. At the same instant, a car about to run her over and kill her (captioned "I'm dead"—accompanied by a thought balloon imagining of her as being dead; she draws "x"'s for her eyes), in front of which she seems paralyzed and unable to avoid the car, even though she knows it is coming (she writes the word "HONK!" above the hood of the car). She appears extremely stressed out (the word "WORRY" is written in dark, bold capital letters) and also resigned to her fate ("Oh well"), and to the expected negative outcome ("I failed").

By contrast, there is no depiction of any external forces beyond her control in the post-intervention drawing; just the sequence of ten images of her proceeding from start to finish through the test-taking process. She shows a variety of appropriate emotions on her face as she goes through the process—from worry (head in her hands) at step 1, through growing nervousness (tap-tap-tap with her pencil) in stage 3, to stress and impatience ("Stop talking and give me the test") as she waits to be handed the test at stage 4, to a big, happy smile for the "A" she anticipates at stage 10. Anticipating her success, at stage 9, she even visualizes rewarding herself with a cookie when the test is over!

The last pair of drawings from the intervention school we describe here shows the pre-post change from a single image of a female student struggling to remain in control and keep a positive attitude to a post-intervention sequence of four images with the student in control, using the HeartMath tools, and feeling successful. In the pre-intervention drawing, the student depicts herself sitting taking the test under pressure (a large spotlight glares down) while she struggles to cope with the stress. In a thought balloon, she tells herself, "No Pressure," "Just breath [*sic*] & relax," and tries to reassure herself, "You know this…", as she holds one hand up to her head in what looks like a gesture of exasperation. By the time of post-intervention measurement, she draws a sequence of four images depicting herself as taking the test and being clearly in control: she even specifically identifies which test she sees herself as taking—"ALGEBRA 1 TEST!!!" She visualizes herself beginning with a highly upbeat and positive attitude (thought balloon: "Test today? Great!"), then uses the TestEdge tools to manage her test anxiety (thought balloon: "HeartMath"), which enables her to then take the test with confidence ("Hey I know this stuff"), and to finish the test with an optimistic feeling ("I think I did fine thanks to heartmath [*sic*])."

These last two pre-post intervention examples are indicative of students who appear to have developed a good sense of emotional self awareness and to have learned to appropriately apply effective emotional management strategies to aid them in the testing situation. These outcomes are consistent with the goal of the TestEdge program.

Pre-Post Examples from the Control School

As indicated at the outset, our examination of all of the pre-post pairs of drawings from students in the two subgroups in the control school revealed fewer instances of positive change. The images in the three pairs of drawings described here represent those that show the clearest evidence of a positive change between the pre- and post-intervention drawings.

The first pair of drawings was classified as a positive change because the images depicted a reduction in an anticipated strong, negative outcome between the pre- and post-intervention periods. In the pre-intervention drawing, the student depicts herself (she has a skirt and long curly hair) as a fully-clothed stick-figure with a big sad face and tears. Her ominous fate—death—awaits her in the background, symbolized by her head (shown detached from her body) hanging, eyes closed (with "x"'s), on a gallows. At the time of post-intervention measurement, she shows only her face, which reveals a lessening in negative feelings and emotions—the tears have gone—although she still appears apprehensive (eyebrows are furrowed), sad (up-turned mouth), and confused (eyes shown out of focus). The image of the gallows—the expected deadly outcome—is gone and no longer seems to be on her mind.

The second pair of drawings shows a change from what appears to be a negative emotion of exasperation and dread—expressed by the underlined word "Ahhh!" accompanied by a tight, guarded smile, suggesting a somewhat positive emotion. The big change is the movement from a single word, with no pictorial representation of the student's image of herself in the pre-intervention drawing, to a full head-and-shoulders self portrait in the post-intervention drawing. In the latter, the lines are strong and confident and show all facial features (open eyes, eyelashes, teeth, dimples, ears), with flowing hair in a bow at the top of her head and a small, heart-shaped necklace around her neck. The overall feeling is one of self-confidence and strength.

The last pair of drawings is from the only student in the control school sample we found in the two subgroups who showed a pre-post change from a single image to that of a sequence of images to represent the experience of taking a test. In the pre-intervention drawing, the student sits at a desk, taking the test while watching a clock on the wall and mechanically saying to himself/herself (in the thought balloon), "Eye on clock, check, check and recheck!," adding "Just do it," as if to assert self will in the face of such stress and uncertainty. By the post-intervention period, the student has expanded this pre-intervention concept into a series of four images with the captions "Do," "Check," "Recheck," and "Turn in." Each of these steps is associated with an emotion shown on a face: a serious expression, a smile, another smile, and a happy, smiling face, respectively. It seems clear that the student is visualizing the test-taking process with himself/herself in control. However, unlike some of the drawings from the students in the intervention school shown above, no emotional management or test stress tools or strategies are in evidence.

Summary and Discussion

A key purpose of collecting student drawings was to generate a more graphic and thereby perhaps more intelligible picture of the impact of high-stakes testing on students. While it is a cliché to say that one picture is worth a thousand words, these pictures convey what many words may not—they provide a window into the inner world of students' thoughts and feelings. Schroeder's assessment was not designed as a clinical psychoanalytical or psychological measure, but one that could be used by non-clinical researchers, school counselors, teachers, and parents. Neither is the goal of the assessment to overly assign meaning to individual drawings, although some may be especially powerful and moving. Rather, the assessment targeted patterns and themes that emerged in the drawings as a whole (Schroeder, 2004: 6).

In this chapter we presented preliminary results from two analyses of the drawings conducted to date—an analysis of a sample of pre-intervention drawings using Schroeder's (2004) Protocol and an analysis of changes in pre–post-intervention pairs of drawings from two matched-group sub-samples. The primary finding from the analysis of the pre-intervention drawings was that an overwhelming majority of the student images conveyed negative feelings and emotions. A second finding of note was that there were few depictions of adults; the portrayals of the few adults who were shown were, almost without exception, negative. More specifically, the following themes were identified in the drawings:

- Depictions of isolation
- Depictions of self-diminishment
- Depictions of fear
- Depictions of extreme anxiety
- Depictions of negativity and emptiness
- Depictions of violence or suicidal ideation
- The use of evocative metaphors, many of which depict darker images, such as students trapped as flies in spider webs, students being engulfed by flames, brains as tornadoes, and so forth.

The analysis of the pre–post-intervention changes in the drawings was conducted on two sub-samples of the intervention school in which we had found a significant reduction in mean test anxiety in conjunction with a significant increase in test performance (White Females in a Regular Class and Math Group 1) and the corresponding matched sub-samples from the control school. From our analysis of the pairs of drawings we identified the following pre-post patterns of change along three dimensions in the intervention school sub-samples:

- Movement from a negative to a positive self-concept.
- Movement from negative to positive feelings and cognitions.
- Movement from negative to positive perceptions of self-control and success.

In contrast to these positive patterns of change, very few instances of positive pre-post changes were observed in the student drawings from the control school sub-samples.

The pre-post pairs of drawings from the intervention school reveal, on the whole, that after participation in the TestEdge program these students appear to have developed a greater degree of emotional maturity with regard to their feelings and attitudes about testing. Some of the images illustrate a clear shift from a strongly negative and unconfident attitude (feelings of anxiety, dread, fear of failure, etc.) to one of confidence and self-efficacy towards test-taking. This likely suggests that students were able to effectively apply the self-management tools and strategies provided in the program. However, even in cases in which a clear movement from a negative to a positive state was not apparent, there were often other types of favorable changes in evidence. For example, one pattern was an apparent growth in students' emotional self-awareness from the pre- to post-intervention drawings. This was conveyed by a greater elaboration of emotional states and by a more candid acknowledgment and representation of students' actual feelings and emotions.

Related to this, the post-intervention drawings also reveal increased objectivity with regard to the testing situation—a greater ability of students to effectively stand outside themselves and see their situation more accurately. This is reflected in the more realistic representation of test-taking as a complex and differentiated multi-step process, as compared to the more simplistic and (often extremely) globalized negative characterization depicted in the pre-intervention images. This was seen in quite a number of the post drawings, where the testing process was broken down into a series of discrete, manageable steps—a portrayal rarely seen in the pre-intervention drawings.

Students' depictions of themselves in their drawings also suggested growth in self-perception. In contrast to the pre-intervention images, in which students often portrayed themselves as barely human stick figures, or in some cases did not show themselves at all, their post-intervention depictions of self, on the whole, were more human-like and realistic, embellished with detail and carefully drawn. This was true even in cases where students still depicted themselves as experiencing stressful or negative feelings.

Although further analysis remains to be completed in order to verify the validity of these preliminary findings, the results are suggestive of a relationship between the TestEdge intervention and the pattern of improvements in students' self-perception, emotional self-awareness, ability to manage test-related anxiety and stress, and feelings of self-efficacy toward test-taking observed in the drawings. Moreover, these preliminary results are also consistent with the broader pattern of findings on the effect of the TestEdge intervention that emerged from the quantitative analyses—namely, that significant reductions in test anxiety, increases in test scores, and improvements in various measures of emotional and social well-being were observed in these intervention school sub-samples.

Overall, the analysis of student drawings provides a richly textured snapshot of students' experience in the current high-stakes testing era. The drawings present compelling evidence of the strong negative emotional reaction that many if not most students have to high-stakes testing—and the likely physiological and psychological costs of such emotional discordance. If such mandated testing is inevitable, it is important for educators to make efforts to mitigate its effects by providing teachers and students with effective tools and techniques to counter the negative, potentially damaging impact of the testing process.

Limitations

Our analysis and interpretation of student drawings had several limitations, as noted by Schroeder et al. (2006). Due to various school- and research-related constraints, the researchers did not have an opportunity to conduct traditional follow-up interviews or discussions with the students to ascertain the accuracy of the meanings attributed by the coders to the images the students had drawn. However, extensive ethnographic observations conducted by the research team and other observers over a six-month period yielded data that appear to corroborate the interpretation of the drawings presented here.

Even so, there are a number of potential drawbacks to using student art in educational research (Schroeder, 2006: 168-9). Viewing art is an inherently subjective experience that often evokes strong feelings and emotional responses, even in researchers who have been trained to dispassionately evaluate an image in relation to a set of explicit, standardized criteria. This is even more the case for art that depicts pain or trauma, and viewers' emotional responses are likely to be further amplified when viewing such depictions created by children. Strong feelings can overwhelm researchers, causing them to lose their objectivity or their ability to accurately perceive what the students are expressing in their drawings (Malchiodi, 1998, p.111). Hence, there is a need for researchers to employ approaches and techniques that will increase awareness of their own personal biases, reactions, and feelings, modulate them, and take them into account when making interpretations of art. Another possible disadvantage is the reticence of some students (particularly adolescents) to express themselves artistically; they may have been socialized to doubt their artistic abilities or to believe that drawing is "childish." Others may view themselves through the internalized lens of a cultural stereotype and thereby render a distorted image of themselves. Yet even these stereotypical images have been found useful by researchers in providing valuable insights into student perceptions.

In sum, the gathering and analysis of student drawings has enabled this study to expand and deepen the view of how mandatory testing is affecting today's students. It is our hope that an increased understanding of this impact will encourage the implementation of programs that can help offset the negative effects that high-stakes testing is having on significant numbers of children and youth.

Chapter XIII
Socioemotional Patterns in the Classroom: Observational Results from the Primary Study

As the primary social unit within the school, the classroom is a rich field of emotions and relations that interconnects the teacher and the entire class, the teacher and each individual student, and each student to his/her classmates. As discussed in Chapter II, the degree to which students are able to learn is strongly influenced by the quality of the socioemotional bonds that connect them to their teachers and their classmates, and to the school community.

Theoretical Considerations

Under optimal conditions, a coupling of positive relations of care and attachment in combination with guidance and firm leadership creates a structure of mutually beneficial cooperative interactions that facilitates and reinforces student learning. However, classrooms in which the relations are largely patterned as negative interactions involving tension, disaffection, or conflict, and in which the relations of guidance are primarily of discipline and control, produce increased stress and anxiety, with poor academic performance a likely result. It is reasonable to expect, therefore, that there would be a relationship between the classroom as a socioemotional environment and student attitudes, learning, behavior, and academic performance.

We expected that the TestEdge intervention would significantly affect the quality of relations in the classroom by improving the socioemotional skills of all classroom members. Therefore, we added an observational component to the study in an effort to map the interactional patterns in the two schools and also to observe teacher implementation of the TestEdge program with their students.

Observational Procedures and Data Collection[24]

To obtain measures of the interactional patterns in the classrooms and, more broadly, within the schools, we constructed an observational protocol (Arguelles & Schroeder, 2004). The protocol was designed to inquire into teaching practices and the relational, affective, and intellectual environment of the classrooms. It included items from instruments that have proven effective in assessing teaching practices, student behaviors, and classroom environments in ethnically and socio-economically diverse schools. School and classroom observations were conducted at three different stages:

1) Prior to the implementation of the TestEdge program, during the month of January 2005 (Time 1),
2) During the implementation of the program in March (Time 2), and
3) After the TestEdge program was completed in May (Time 3).

The Time 1 and Time 3 observations were used for the analysis of potential differences that could be attributed to the TestEdge program and the Time 2 observations were used primarily to provide a qualitative perspective on the fidelity of the implementation.

Doctoral students from Claremont Graduate University's School of Educational Studies, who had research expertise and were familiar with the TestEdge program, conducted the observational research. In three separate sessions, CGU observers were trained by Professor Arguelles and Dr. Bradley in observation methods and calibrated in relation to one another to ensure operational consistency in the application of qualitative data-gathering procedures. Observers also conducted observations during a pilot

[24]This section draws heavily on Teri Hollingsworth's work (Hollingsworth, 2006)—a Claremont Graduate University graduate student member of the TENDS project.

study to practice skills for collection of data. The role of observers at both sites was that of a passive observer only. There was no participation in classroom activities and no interaction between observers and students; both the teachers and the students were informed of the observers' passive role.

For consistency among observers, we constructed a "Classroom Observation Form" and an "Observational Protocol" (Arguelles & Schroeder, 2004) to provide operational guidelines for recording classroom observations on 70 variables organized into a conceptual framework of eleven categories.[25] The observation for a given variable was rated on a 5-point Likert scale. Sections of the form also provided specific detailed instructions about the meaning and relationship of measured variables. In an effort to ensure accurate recording of the observations, observers were instructed to complete the form directly upon exiting a classroom and prior to conducting another observation.

The uniform framework of the Observational Protocol provided a means for standardized measurement enabling systematic comparison of observational data collected from all classrooms. Each observer used the form in training and was aware of its contents during observations; the form gave the study a systematic vocabulary with which to record a standardized set of variables in seventeen different 10th grade high school classrooms. However, as Hollingsworth (2006) rightly notes, a limitation of the predetermined nature of the variables in the Observational Protocol is that the coded data may not always reflect the specific variables that individual observers saw as significant in the classrooms they observed. In order to ensure that observation was not just confined to the dimensions of the protocol, each observer was also trained and instructed to record other behaviors and events they observed as well in a field journal.

Two designated observers were assigned to each primary school site. Observers visited each site for a full school week in January (Time 1), March (Time 2), and May (Time 3), 2005. Because of the scope of the project, it was not possible to have an observer in each class. One of each of the eight English teachers' classes at the intervention site was randomly assigned to the two observers. These classes were observed twice on each of the three visits, six times in all. The same process was used at the control site, where two observers were randomly assigned to observe classes of nine

[25]These instruments can be viewed on the Institute of HeartMath's website, www.heartmath.org.

different teachers. One difference was that at the control site observations were conducted in 10th grade Social Studies rather than in English classes. This was not considered a problem because no curriculum intervention was introduced at the control site. During the three observational periods, each of the assigned classrooms was observed twice during the week, for a total of three-and-a-half hours at the intervention site and one-and-a-half hours at the control site. The difference in observation time was a result of the intervention site's block schedule that made class periods twice as long (90 as compared to 45 minutes).

Before describing the coding procedures and results of the quantitative analysis of the observational data, it is helpful to provide some descriptions of the physical and socioemotional context of the classrooms and the interaction patterns observed between teachers and students. Readers should be aware that field observation notes included below were written in the moment and are not polished in terms of style and grammar. They also invariably reflect the individual personalities and subjective proclivities of the observers. Nevertheless, they are part of the observational record and provide a first-hand view of the phenomenological realities of the school and classroom environments in which the study was conducted.

Physical and Social Context[26]

As described above in Chapter V, the intervention school is an older campus surrounded by a predominantly Hispanic community. It is a large school of approximately 3,000 students, the majority of whom are first and second generation Hispanic from middle and low-income families. By comparison, the control site is a newer, smaller school of approximately 2,000 students populated primarily by White students from a predominantly middle and upper-middle income community. The teacher workload was higher and the per pupil expenditure lower at the intervention site, compared to the control site. There were also differences between the two schools in educational goals. The intervention site school's goals focused primarily on creating a safe, friendly school environment. The control site school emphasized preparing students for college and university admission. These differences were reflected in observations of the social ecology of the two campuses by the CGU research team.

[26] This section and the following section, "Classroom Interaction," draws heavily from the CGU team's report and the field notes of their observations of the two schools (Arguelles et al., 2006).

Intervention School Site

Impressions of the intervention school campus are found in the following observer notes:

> The gates along the fences are kept locked after 8:30 am until school lets out for the day, presumably protecting the students and campus from the outside and keeping the kids where they are supposed to be.
>
> I saw that security-personnel were walking around gathering trash and monitoring students who were not in class. While students were busy in class, the large security staffs kept the environment not only secure, but also clean which was another aspect of safe[ty]. I observed no graffiti on the campus. I also observed what looked to me like security personnel checking student lockers.
>
> At lunch the campus was quiet and peaceful, students were grouped mostly by ethnicity, but I felt no tension between the groups. There was an incident at lunch the first day … walking across the entire campus … I notice that a large crowd of students (had gathered).… I looked for the security personnel, but could not find them. It looked as if there were a thousand students in one small area.… I later heard that I had just missed a fight.

The majority of students at the intervention site were Hispanic and Caucasian; there were few African Americans, and even fewer Asian/Pacific Islanders. The typical attire of male students was sweatshirts that usually came with a hood (referred to by students as a “hoodie”) and, in some cases, designer sneakers. Those not wearing sweatshirts usually wore a sports jersey or a white extra-long t-shirt. Female students displayed more variety—short tops, tight jeans, pony tails, and hoodie jackets were the norm. Small purses that matched the outfit of the day were typical. Some students wore Navy ROTC uniforms periodically.

On the last visit, in May, many male students were described as wearing jeans that hung below their buttocks. There were pockets of students who seemed to dress and wear their hair in a similar fashion, likely signifying that they belonged to a group with a dress code. Some students wore athletic-type gear in the school team colors. On one instance, observers noted a group of Hispanic students who were all wearing black and standing around in a circle, with one student standing in the center of the circle. Each student on the outside came up to the student in the center and formally shook his hand, as if participating in a group ritual or ceremony.

Control School Site

Although there were differences in student attire between the two schools, with students at the control site featuring more affluent clothing, there were also similarities, as the following note suggests:

> Students are dressed in a youthful, but affluent manner. The prevalent style, both for males and females, was to wear hooded sweatshirts. Most wore the hood on their shoulders, but a few pulled the hood up over their heads.

On the first visit to the control site, observers saw students who seem cohesive and supportive of one another. An important mention was made of the social cohesion between students—how much the students seemed to like and respect one another.

On the second visit, however, when mandatory testing was imminent, observers noted a change in mood:

> The students in general seemed off [balance] this time. The reason was not known to me at the time I did my notes, but I [learned] later of two major happenings. Two weeks prior, there had been a bomb threat and all the students were evacuated. Then one week later a popular senior girl committed suicide. Mandatory tests were coming up, so when you add these elements together you get a tense and bleak school environment.

Classroom Interaction

In both schools a wide variance in classroom social climate and interaction was observed, shaped, in large part, by the unique combination of the teacher and group of students involved. In most classrooms, the socioemotional tone and interactional structure was set by the personality and teaching style of the teacher. However, in a few classes in both schools, the students were observed to strongly influence the dynamics and mood of the classroom.

To provide the reader with a sense of the socioemotional tone and an image of the interactional patterns, we present observer descriptions of classrooms from each school.

Intervention School

Classroom A

The room is a little dark, with no direct sunlight. The lighting is fluorescent. It is a little stuffy in the room. The overall effect is cramped and cluttered, but not chaotic.

There is a lot of interaction between the students and the teacher (a young athletic-looking female), as the students come in and get settled. The interplay is mutual; she initiates contact as well as responding to students, and they do the same. She moves through the entire room as she explains the assignment. She personally greeted each student during this period.

The students and the teacher are comfortable with each other; there is a nice energy flow in the room. It feels like a safe place. The teacher seems to really enjoy the kids, and the feeling is mutual. Her posture is open: she does not sit behind a desk, she puts her hand on students' shoulders, she makes good eye contact. The students, although in a variety of postures, are relaxed and present. When the students work on the assignment, the energy is quiet and focused in the room.

Classroom B

The teacher is a tall, heavy-set woman wearing a heavy key chain hanging from the front pocket of her pants. Her range of facial expressions went from serious to irritated; she did not smile even once during the observers' first visit.

On another observation, the teacher looked a little surprised to see the TENDS team. She also did not seem too happy about [what she experienced as] the intrusion of observers in her class. There were student names along the right hand side of the blackboard, along with their disciplinary infractions: C=cursing, I=inappropriate comments, B=behavior, and D=defiance (attitude).

The teacher used the overhead projector for notes about tragic heroes, and Shakespearean tragedy. The students were directed to copy them. There was no interchange of any kind, no engagement, all rote behaviorist methodology; a very teacher-focused class. Next, she gave the students an open book quiz for Act I of Julius Caesar. There was very little interaction between students, other than furtive interchanges. There was very little student-initiated contact with the teacher. The interactions from teacher to student are limited, and mostly punitive. After students left, the teacher stated that she did not believe that the students will be helped by the *TestEdge* tools.

Classroom C

The teacher, a middle aged female, veteran teacher asked her students to open the window blinds to, "help her flowers grow, because they needed sunshine." She was referring to her students. Her methodology was more than simple constructivist: she had formed a caring relationship with her students that held them responsible, yet supported them at the same time.

The energy flow shifted naturally from one student to another. In the back of the room there was lots of whispering, lots of banter and quite a bit of free flow talk back and forth between the teacher and the students. There was much talk of college. The teacher seemed to have a clear line between two goals—test prep work, and work for intellectual goals such as college.

Students moved to group projects and it was obvious that the classroom was student centered at this time. Thus, the teacher versus student centeredness, seemed to move back and forth in this classroom. There was more evidence of the "student centeredness" of the classroom during the group activities over the two observations. When the teacher asked students to work on their projects, students were all over the classroom working on numerous missions. Many students worked in groups, many on their own, many changing roles during this time.

In terms of themes in the field notes, observers recorded a wide range of classroom interaction patterns. There were instances of sleeping in class, coming late to class, talking back to the teacher, and engaging in non-class-related activity. There were a number of classrooms in which the students appeared to be just passing time—unengaged and bored. There were also were instances of rowdy, disruptive, and disrespectful behavior toward other students or the teacher which, in a few cases, led to disciplinary action. But there were also classrooms in which students were focused on learning, were actively engaged and enthusiastic, and in which highly respectful and cooperative patterns of interaction were observed.

Another pattern involved an association between styles of student attire and the emotional mood of some classrooms. In two open and interactive classrooms, students seemed to wear less black and keep their hoods down on their shoulders. Girls wore ponytails and light colored sweaters. In two other classrooms students wore dark colors which added to the dark feeling in the room itself. Several students had their hoods up over their heads so far that their faces were not visible.

Control School

Classroom D

From the beginning the class was very rowdy. The teacher, a warm, gentle man, began giving instructions very loud to overcome the noise of the class's talking and playing. At least one-fourth of the class was talking with each other and acting as if no teacher was at the front of the room. He seemed unconcerned that there was so much commotion and just kept addressing the class as the talking continued.

Trying to hold the attention of the class, the teacher continued to discuss the intellectual content (history) of the class and most students calmed down. Two students laid their heads on their desks. There was a lot of wiggling, twitching and shuffling going on amongst the students. The energy in the room was very high but not totally out of control. The teacher turned on his laptop and LCD projector and began to type his outline as he talked about the era of history they were covering. Students were laughing and talking all over the room as he instructed.

Some students were paying close attention to the teacher. Many others were taking notes, but no one seemed to be bothered by all the talking and laughing. The teacher positioned himself at the front and center, not moving, just lecturing and typing. He showed kindness and respect for the students in their interactions and they seemed to take full advantage of it.

Classroom E

One boy threw a ball of paper across the room aimlessly. The teacher (a male) was quick with a firm reprimand that let the student know he was not in trouble, but there was also a clear understanding that he was not to do it again. The teacher called for attention and everyone immediately responded. The lesson was to be about test taking. As the teacher began to ask questions, the students were quick to respond and the teacher recognized them immediately.

As the teacher lectured, students continued to have rapt attention. He began asking specific questions to specific students. Usually two to ten

> of the seventeen students present would respond to the questions. The class atmosphere was warm. The teacher began to call on students who weren't participating, by name, and the response was always quick. He affirmed each student's response no matter what it was. He called on several students who seemed to be confused. If they stammered, he was kind and said that he would get back to them and did. The energy level of the students increased as they were affirmed; they seemed very pleased with themselves.
>
> Some students started mumbling and the teacher halted it immediately. He was quick to correct behavior of any student, but did so in a caring but firm manner. Students seemed to show respect for his commanding presence. There were lots of smiles and laughter in the room.

Based on the data recorded in the field notes, observers saw only a few moments where issues with behavioral problems occurred. Observers noted students sleeping in class, coming late to class, talking back to the teacher, and cell-phone use, but mostly students were focused on learning.

Observation of the TestEdge Intervention

As described earlier, the implementation of the TestEdge program took place in three phases. The first phase was a one-day off-campus Resilient Educator workshop, facilitated by senior HeartMath trainers, in which the teachers learned the core tools and techniques of the HeartMath System and how to apply them in their own lives. After a three-month period, intended to give teachers ample time to practice and internalize the tools, the teachers attended a second one-day workshop that specifically focused on the content of the TestEdge program and best practices for teaching the program to high school students. Finally, the third phase involved the teachers conducting in-class instruction of the TestEdge program throughout the Spring semester. This included classroom time focused on the key concepts and techniques presented in the program as well as time in the computer lab to work with the Freeze-Framer heart rhythm coherence feedback system.

TestEdge Instruction

Some teachers were very enthusiastic about the TestEdge program and were committed to making the intervention successful, as can be seen in the observer's field notes on one teacher's instruction of the program:

> After the journal activity, students worked in their *TestEdge* text. They were on Part Two of Lesson Six, attitude breathing. To introduce the lesson the teacher explained "attitude breathing." First he asked them to remember an attitude. Several students raised their hands. He asked how many had practiced. How many noticed a change? Several students had comments about practicing.
>
> He then asked students to put their hands on their chests so that they would "look as foolish as he did," and asked them to focus on breathing slowly. He used humor while he demonstrated this. Students followed along. All were introspective. When asked to think of a positive feeling, many had big grins on their faces. He continued this practice for about two to three minutes.
>
> He asked if anyone had anything to share. Some students agreed that it made them feel light. One said it made her feel more relaxed. After each comment, the teacher acknowledged the students comments, and then added his own insight. Then he said, "You guys need to practice this: the more the better." He then had them read along in the book and asked them to fill in the workbook. As they worked, it was almost completely quiet in the classroom. As they left he asked the students to "please practice before the next class."

Other teachers found it more difficult to facilitate student interest with the content of the program and to engage them in practicing the TestEdge tools, as can be seen from the observer's field notes on another teacher:

> Students pass out the *TestEdge* booklets and the teacher reviews what the students had previously learned. She asks them to practice "The Neutral Zone": "Self control. Hand on your heart," she instructs as she waits for them to quiet down. A student asks, "Do I have to?" "Yes," she says. "Close your eyes if you want to. No talking. Focus on your heart area. Breathe four seconds in, four seconds out." She sat on a stool in front of the class, hand on her heart, modeling the tool for students. It became quiet and restful in the classroom.
>
> Further along, she brings up the topic of the CAHSEE and reminds the students that how they perform on this exit exam will determine whether they graduate or not. The anxiety level seems to rise a lot as the students ask many questions.
>
> She then moves on to "Attitude Breathing," pointing out the Attitude poster as a guide: "Hands on your heart if you want…practice your breathing… deep breaths…focus on the area around your heart…" The room was very quiet, the students sat with their hands on their hearts along with the

> teacher. There appears to be no discernible anxiety in the room. But not all of the students are involved as some are fidgeting and looking around the room. "STOP LOOKING AROUND AND DO IT!" she commands.
>
> After the exercise was over, the teacher asked the students to share the attitudes that they want to change. There were no answers. She tries to get a response by sharing her own attitudes. There were still no answers.

A small number of teachers were highly skeptical or even resistant to the TestEdge program, as the following excerpt from the observational notes of another teacher's class shows:

> The teacher began the class with HeartMath. She started with breathing practice, and asked students to "Mellow out, take it down a notch with breathing," while she took attendance. I noticed that the TestEdge Neutral Chart was hanging on the front of the blackboard.
>
> The students had the *TestEdge* booklets out and were working on Chapter 6. The teacher seemed a little nervous as she read the chapter out aloud. She sighed audibly several times as she went along. The students were watching her carefully. A male student asked if he could read a quote from the booklet—it was a quote about roses and thorns. The teacher exclaimed "Ooohhh!" sarcastically. The boy persevered by talking about how similar the metaphor was to the glass half empty/half full metaphor.
>
> The teacher continued to present the material in a perfunctory way that had a clear undertone of skepticism. At one point she made reference to "TestEdge drama." The class's work on Lesson 6 ended without the teacher giving the students an opportunity to practice any of the TestEdge tools.

This was the same teacher who had previously told an observer that she did not believe the students would be helped by the TestEdge tools.

Freeze-Framer Training

An important element of implementing the TestEdge intervention involved teaching students how to use HeartMath's computerized heart rhythm coherence-building technology (Freeze-Framer system). The teacher and students of a given class went to the computer lab, where 33 computers were set up on tables. Each student sat in front of a computer monitor and attached a finger or ear lobe heart rhythm sensor. The computer monitor displayed real-time data showing the student's heart rate and heart rhythm activity and ratios of low, medium, and high heart rhythm coherence. Once the student learned how to use the system and could sustain a medium level of coherence, s/he was allowed to proceed to three different computer games driven by the level of coherence s/he could maintain. What follows is a description from an observer's field notes of the Freeze-Framer training session held in a computer lab for Classroom A's students:

> The teacher stood at the door of the lab, welcoming students as they came in, and directing them to computers. The mood in the room was upbeat; the students were clearly happy to be using the computers today. There was lots of chatter, and it felt a little chaotic as the HeartMath team prepared to start. Jackie (the HeartMath trainer) began by getting the attention of the students, with some help from the teacher, who was maintaining a strong presence in the room. The students were squirrelly, but they were settling down pretty quickly. Rollin and Del (also from HeartMath) … stood amongst the kids, patrolling a bit; the students were very aware of them, sitting up straighter, focusing on Jackie more as they walked by.
>
> Rollin explained the finger sensor to the class; Jackie, Del, and the teacher helped the students get it on right. There was a lot of talking amongst students and the adults kept the students on track while not squashing their enthusiasm. The students appeared eager to move on and begin using the Freeze-Framer; Jackie was going over the preliminary information methodically and was quick to ask students to stay with her.
>
> As students saw their pulse and heart rates on the screen for the first time, they were excited. There was laughter and many comments: "Dude, yours is RAD!" "Why is mine going so fast?" "Why is there red on mine?" The energy level among the students was very high and not contained at this point. There was an underlying anxiety, as the students were not clear on what this meant yet, and what they would do next. All four of the adults were needed to keep the anxiety from escalating by giving students information and assistance on an individual basis.

> The teacher explained the concept of coherence to her students; she had been watching Jackie and her kids closely, and jumped in to make sure they were getting it. She told them: "Focus, no wiggling, concentrate, and don't talk to your neighbor." She seemed to know the TestEdge and Freeze-Framer material and related it to what she had been teaching (TestEdge tools) in the classroom for the last several weeks. She continued "If you want to get green, you have some work to do!" (The students had to achieve a score of over 50% at the blue or green levels [medium and high coherence, respectively], in order to play the Freeze-Framer games).
>
> The students are now doing 5 minutes of breathing: some with heads down; some with sweatshirt hoods pulled up and their eyes; some had legs bouncing; others let out sighs and groans of frustration. The adults address the overt stress responses.
>
> The students who reached a high enough level of coherence were allowed to begin playing the "Meadow Game." The teacher sat at a computer left vacant by a student called in to the office. She modeled by participating in the Freeze-Framer session herself. A student who had been identified as a "bully" pulled his hood down over his eyes; when he sat that way for a minute or so, his body posture shifted from agitated and fidgeting to relaxed and calm—a dramatic change. Playing the "Rainbow Game," he now had a full of pot gold coins at the end of the rainbow, indicating that he is coherent.
>
> Jackie wrapped it up, saying "Good-bye," as they began to move en masse out of the room. One male student called out, "Thank you HeartMath team! I had a great time!"

Patterns of Change in Classrooms: Quantitative Analysis

Following the completion of observations, Classroom Observation Forms were collected from each observer, entered into an SPSS database, and descriptive statistics—pre and post—were computed for each variable, including frequency of observation, mean and S.D. of observer rating score. In her analysis of the classroom observation data, Hollingsworth (2006) used the pre-mean and post-mean score for each variable to compute a measure of the Time 1 to Time 3 difference in means. The results of her analysis are shown in Table XII.1. In the table the direction of pre-post change is indicated by the following notation: "P" signifies positive change (supportive of student learning); "=" signifies neutral (no change); and N signifies negative change (not supportive of student learning).

Table XII.1 Analysis of Pre-Post Classroom Observation Changes: Mean Observer Rating Score by Intervention Status

	Intervention School		Control School	
	Change in Mean Score January–May	Direction of Change	Change in Mean Score January–May	Direction of Change
Condition of Classroom				
Crowding	-1.25	P	-0.06	P
Clean	-0.38	N	0.13	P
Adequate Lighting	-0.13	N	0.01	P
Adequate Ventilation	-0.63	N	0.01	P
Comfortable Temperature	-0.5	N	0.13	P
Orderly/Organized	-0.5	N	-0.28	N
Safe	-0.63	N	-0.26	N
Pleasant Ambiance/Atmosphere	0	=	-0.39	N
Affective Mood in Classroom				
Happiness	0.25	P	-0.68	N
Sadness	0.13	N	0.56	N
Anger	0.13	N	0.5	N
Fear	-0.25	P	0.29	N
Frustration	-0.38	N	0.43	N
Excitement/Engagement	0.13	P	-0.47	N
Anxiety	0.13	N	0.06	N
Motivation	0	=	-0.63	N
Listless/Lethargic	0.63	N	0.37	N
Peaceful/Serene	0.25	P	0	=
Energetic	-0.13	N	-0.56	N
Boredom	0.38	N	0.2	N
Intellectual Behaviors				
Persistence	0.25	P	-0.77	N
Impulsivity	-0.25	P	-0.12	P
Empathic Listening and Understanding	0.63	P	-0.7	N
Questioning/Problem Solving	-0.5	N	-0.46	N
Creativity/Insightfulness	-0.13	N	-0.32	N
Humor	0.13	P	-0.37	N
Curiosity/Wonderment	-0.13	N	-1.01	N
Cooperative Thinking (Social Intelligence)	0.25	P	-0.18	N
Cohesion				
Emotional Bonding	0.13	P	-0.78	N
Supportiveness	-0.06	N	-0.75	N
Boundaries	0.13	P	-0.52	N
Receptive Student Interactions				
Accept Compliments/Positive Feedback from Others	0.11	P	0.26	P
Appropriate Response to Other-Student-Initiated Contact	-0.13	N	0	=
Expressive Student Interactions				
Initiate Contact with Teacher	-0.25	N	0.38	P
Ask Appropriate Questions	-0.5	N	0.26	P
Initiate Contact with Other Students	0.38	P	-0.13	N
Cooperate with Other Students	0.5	P	-0.04	N
Cooperate with Teacher	-0.25	N	-0.51	N
Effective Cooperative Group Work	-0.05	N	1.07	P
Collaborate in Peer Learning Groups	-0.05	N	0.73	P
Respond Appropriately to Teacher Verbally	-0.5	N	-0.32	N
Respond Appropriately to Teacher Nonverbally	0	=	-0.4	N
Sociometric Measures				
Peer Acceptance of Students	-0.13	N	-0.2	N
Peer Rejection of Students	0.38	N	0.14	N
Peer Neglect of Students	0.13	N	0.06	N
Instructor Acceptance of Students	0	=	-0.67	N
Instructor Rejection of Students	0.63	N	-0.11	P
Instructor Neglect of Students	0.75	N	-0.46	P
Prosocial Behaviors				
Assertiveness	0.13	P	-0.38	N
Leadership	0	=	0.1	P
Respect	-0.5	N	-0.4	N
Collaboration	0.18	P	0.46	P
Cooperation	0	=	-0.12	N
Engagement	0	=	-0.5	N
Empathy	0	=	-0.78	N
Generosity	-0.13	N	-0.87	N

Direction of Change
P Positive change
N Negative change
= No change

Table continued on next page

Continued: Table XII.1 Analysis of Pre-Post Classroom Observation Changes: Mean Observer Rating Score by Intervention Status

	Intervention School		Control School	
	Change in Mean Score January–May	Direction of Change	Change in Mean Score January–May	Direction of Change
Antisocial Behaviors				
Hostility	0	=	0.08	N
Aggression	0.5	N	0.09	N
Withdrawal/Disengagement	0	=	0.51	N
Flexibility				
Leadership	0.23	P	-0.09	N
Discipline	0.19	P	-0.14	N
Negotiation	0.5	P	-0.2	N
Roles	0.31	P	-0.08	N
Rules	0.31	P	0.02	P
Communication				
Listener's Skills	-0.13	N	-0.79	N
Speaker's Skills	0.07	P	-0.61	N
Self-Disclosure	1.3	P	-0.5	N
Clarity	-0.79	N	-0.49	N
Continuity/Tracing	-0.38	N	-0.63	N
Respect/Regard	-0.75	N	-1	N

Direction of Change
P Positive change
N Negative change
= No change

Pre-Post Results[27]

Before discussing the results, a few caveats are in order. Due to the small case counts of observers (N = 4, 2 in each school) and observations (N = 24, 2 classes x 6 times x 2 schools), we did not compute the statistical significance of pre-post change measured on the variables. For these same reasons, we also did not conduct a validity and reliability of measurement analysis on the set of variables within each category of the observational protocol, which would be necessary to treat each formally as a statistical construct. Thus, in interpreting the results shown in Table XII.1, we view any patterns of pre-post change as suggestive rather than definitive. In what follows, the results are organized by the major category of the observational protocol.

Condition of Classroom

At baseline (Time 1), with the exception of Crowding, which was rated "high," the physical conditions of the classrooms in both schools on the other seven variables were rated "good" to "excellent." However, by Time 3 in the experimental school, a negative change in classroom conditions was observed across all variables, except for Crowding (improvement) and Pleasant Atmosphere (no change). This pattern was in contrast to that in the control school, where a positive change in classroom conditions

[27] This section draws heavily from Hollingsworth (2006).

was observed on five of the eight variables—Crowding, Clean, Adequate Lighting, Adequate Ventilation, and Comfortable Temperature.

Affective Mood

The baseline observations of classroom affective mood showed that while students in both schools were Happy, Excited/Engaged, and Energetic, those in the control school classroom were rated with higher mean scores on these indicators. However, after the intervention by Time 3, whereas there is evidence of some improvement on four of the twelve measures of affective mood in the experimental school, in the control school a decrease is apparent on all measures, with one exception—Peaceful/Serene, which remained constant. Specifically, in the intervention school, improvement is apparent for Happiness, Fear, Peaceful/Serene, and Excitement /Engagement, while a slight negative change is evident for Sadness, Anger, Frustration, Anxiety, Listless/Lethargic, Energetic, and Boredom.

Intellectual Behaviors

At baseline, whereas none of the eight variables measuring various aspects of intellectual behavior were observed frequently in the classrooms of the intervention school, four aspects of intellectual behavior were frequently seen in the control school—Empathic Listening, Questioning Problems, Humor, and Curiosity/Wonderment. However, while a decline on all aspects, except Impulsivity, was observed in the control school, an improvement on five aspects of intellectual behavior was evident in the intervention school classrooms—Persistence, Impulsivity, Empathetic Listening and Understanding, Humor, and Cooperative Thinking.

Cohesion

While there was some variation by classroom in the intervention school on the three measures of Cohesion at Time 1, the means for the variables indicated an overall low level of cohesion (corresponding to the "disengaged" or "separated" categories). This was in contrast to the baseline pattern observed at the control school where the means for Emotional Bonding, Boundaries, and Supportiveness were indicative of a more socially connected, cohesive group of students in the classrooms. However, at the time of the post-intervention observations in Time 3, whereas there was a decline on all three measures of cohesion in the control school, an improvement on Emotional Bonding and Boundaries was observed in the intervention school; this was accompanied by a slight negative change for Supportiveness.

Receptive Interactions

In relation to the two measures of student relational receptivity—Accept Compliments/Positive Feedback from Others and Appropriate Response to Other Students—observers recorded seeing these behaviors only "occasionally" in the classrooms in both sites at Time 1. By Time 3, improvement was observed in student receptivity to compliments and positive feedback from their peers in the classrooms in both schools. However, while no change was observed in the control school on Appropriate Response to Other Students, a slight decrease in this relational response was seen in the intervention school's classrooms.

Expressive Interactions

Of the nine measures of student interactional expressiveness, two—Effective Cooperative Group Work and Collaborative Peer Learning Groups—were not observed at all at baseline in the intervention school classrooms. The other seven aspects of expressive interactions were seen only "occasionally." By contrast, at the control school, while Initiate Contact with Teacher and Cooperate with Other Students were observed "occasionally," Ask Appropriate Questions, Initiate Contact with Other Students, Cooperate with the Teacher, and Respond Appropriately to the Teacher Verbally (/Nonverbally) were "frequently" observed. In the intervention school classrooms by Time 3, while some improvement was evident on Initiate Contact with Other Students and Cooperate with Other Students, a negative change was recorded on all of the other measures, except for Respond Appropriately to Teacher Nonverbally, which remained unchanged. By comparison, classrooms in the control showed an improvement in four measures of relational expressiveness—Initiate Contact with Teacher, Ask Appropriate Questions, Effective Cooperative Group Work, and Collaborate in Peer Learning Groups—and a negative change in the remaining five measures.

Sociometric Measures

Three aspects of the quality of relations (acceptance, rejection, and neglect) were observed and recorded for interactions among students and between the teacher and students. In classrooms at both schools at baseline observation, there was little evidence of rejection or neglect by teachers or peers. Rather, students and teachers at both schools were observed to be accepting of their peers and accepting of their students, respectively. However, by Time 3, with the exception of Instructor Acceptance of Students (which remained unchanged), all sociometric indices declined in classrooms at the intervention school. And while a similar pattern of decline in most of these indices was observed at the control school, some improvement was observed in how teachers related to their students on two measures—Instructor Acceptance of Students and Instructor Neglect of Students.

Prosocial Behaviors

As Hollingsworth notes, when making observations in the prosocial behaviors category in classrooms with students from diverse social backgrounds, field workers were cautioned to be sensitive to the impact of culture, ethnicity, gender and family systems norms on student behavior. Of the eight measures of prosocial behaviors, only Respect was observed frequently in the intervention school and Assertiveness, Leadership, and Engagement were seen occasionally. The four indicators of relational mutuality—Collaboration, Cooperation, Empathy, and Generosity—were recorded infrequently. In the control school classrooms, by contrast, there was greater evidence of relational mutuality in that, along with Respect, Engagement, Empathy, and Generosity were also observed frequently. Leadership, Collaboration, and Cooperation were observed "occasionally." In terms of change by Time 3, improvements were evident for Assertiveness and Collaboration at the intervention school, while Respect and Generosity had declined somewhat. At the control school, by comparison, positive gains were observed for Leadership and Collaboration, while the other six measures of prosocial behaviors showed a negative change.

Antisocial Behaviors

In both schools at baseline, there was little evidence of antisocial behavior recorded in any of the three measures—Hostility, Aggression, and Withdrawal/Disengagement. By Time 3, while all three measures showed a degree of negative change in classrooms at the control school, only Aggression had worsened slightly in the intervention school; Hostility and Withdrawal/Disengagement remained unchanged.

Flexibility

Within the context of today's multicultural schools, the degree to which the social order of the classroom is flexible—adaptive to the sociocultural diversity of the students—was measured on five significant sociological dimensions: Leadership, Discipline, Negotiation, Roles, and Rules. At the time of baseline measurement, the observations indicate that whereas the social order in the intervention school's classrooms involved more social control—all variables were rated from "structured" to "rigid"— in the control school, with the exception of Leadership (rated as "structured), the classroom order was more flexible and socially adaptive. Interestingly, by Time 3, across all five measures there was evidence of a consistent change toward a more flexible social order in the intervention school. By contrast, with the exception of a slight improvement in Rules, all other measures showed a slight a negative change in classrooms at the control site.

Communication

In both schools at baseline, the six measures of communication in the classroom were rated high. However, by Time 3 a consistent pattern of negative change was observed across all measures in the control school. Although this pattern of negative change was evident for four aspects of classroom communication in the intervention school—Listener's Skills, Clarity, Continuity/Tracing, and Respect/Regard—an improvement was observed on two measures—Speaker's Skills and Self-Disclosure.

Summary of the Observational Findings

The aggregated results from the analysis of observational data suggest that when compared to the control school, more positive changes were observed at the intervention school—on 24 of the 70 variables (34.3%), versus 16 of 70 (22.9%) at the control school. Even more striking is the difference between the schools on negative changes, in which many more variables were observed to move in a negative direction in the control school—62 of 70 variables (88.6%), compared to 36 of 70 (51.4%) in the intervention school. Ten (14.3%) of the variables remained unchanged at the experimental site whereas 2 (2.9%) remained constant at the control site. Overall, on 48.6% of the variables either positive change or no change was observed in the intervention school, whereas this was true for only 25.7% of the variables in the control school.

Hollingsworth (2006) points to several patterns observed in the qualitative data collected from the intervention site:

- Overcrowding and clutter in many classrooms;
- Extreme negative to extreme positive student affect/mood in classroom environments;
- Extreme variability among classrooms in levels of affective energy expressed by teachers;
- Fluctuating patterns of student responsiveness in the various classrooms;
- High levels of student disengagement across many regular education classrooms;
- Absence of leadership roles among students; and
- Lack of antisocial behavior among students.

The following patterns were evident in the classroom observational data collected at the control school:

- General student focus on learning;
- Variety of teaching styles; and
- Different levels of teacher influence.

Post-Intervention Teacher Interviews[28]

In the teacher interviews, conducted in the two-week period following the Time 3 observations, one of the questions asked teachers whether they felt "adequately prepared," prior to the TestEdge program, to help students manage fear and stress in relation to test-taking. Many of the teachers indicated that they had no specific background or training to assist students in this regard, other than "common-sense approaches":

> "I never had an actual tool or new concept to help with their fear or anxiety about tests."

> "No. I felt like I'd done everything this first year as baptism by fire. I taught them how to study."

[28]As with verbatim material from field worker notes presented earlier in this chapter, we have quoted from the interviews without editing.

"I was not as well prepared. … The TestEdge tools provide me with a readily accessible method that addresses (my students') emotional state, their physical sensations and intelligence at the same time."

A number of teachers reported that some students had become very engaged in using the TestEdge tools to facilitate a positive change in their attitudes and behaviors. One teacher said she believed the tools would have a positive effect on her students' future performance, and another expressed confidence that her students would carry the tools they learned in the program with them after they graduate. Teachers also reported observing students making more of a connection between how they feel and how they perform:

"I noticed they would really become involved in shifting their focus and attitudes. There was a change in attitude for some students. I would think that positive attitudes and remaining calm would help their academic performance."

"A lot of students doing poorly in class liked the TestEdge course."

"The Honors kids took to it like ducks to water—managing stress and improving areas of stress performance."

"One student was going through the worst year of his life (family problems); he's passing now."

"[One of my students] said tests freaked her out. I gave her one of the *TestEdge* books I got. She went to Mexico for one month and I told her to read it. … She found the tools helpful beyond class. It helped her with rude customers at her restaurant job waiting tables."

Aside from the reported successes with the students, many teachers reported positive experiences in various situations from their personal practice with the tools:

"I use the HeartMath tools all the time. I have a two- and a five-year-old. My mother lives with me. I teach three preps. I definitely find the tools useful. They also help me with my own school work; I'm getting A's in graduate school."

"Yeah, it helps in traffic. It helps me maintain my composure when my students are rowdy. It helps me slow down and not react when dealing with an upset parent. I take a moment to maintain a positive mind frame when facing angry situations between two students, between students and teacher, between teachers and their own uncertainty and not get

> caught up in how the other feels. I can get an easier and wiser solution to what I'm facing. It's very effective in not allowing stress of others to become my own."
>
> "I also was crazy enough to take on STAR Testing Coordinator. … It was a nightmare, logistically. I would have been sick if I hadn't had HeartMath … and appreciate that I'm still alive or haven't had a heart attack."

However, there were also teachers who were more skeptical or cynical about the TestEdge program, as exemplified in the interview with this teacher:

> "There is a presumption that teachers need to help students manage fear and stress about taking tests. My main job is to help students become citizens of our country. Best I can do is transfer knowledge, then they need to discover things on their own. I'm not a counselor or psychologist so I don't think my job is to teach students to manage stress. It's a bit intrusive."
>
> He doesn't agree with the basic premise of TestEdge program that we should be helping students transform emotions or change neural circuits.
> "That goes against the idea that students should have natural emotions. Some of our most creative artists, geniuses, etc. painted out of angst. Some of our best artwork came out of unhealthy emotions." To "suppress emotions" goes against what he believes.
>
> He had a hard time buying into the initial Resilient Educator training. You had to have a belief that the principles were right and bought into it prior to going to the training. It was "too vague and touchy-feely." … He was also not happy with how much class time this took.

Time 2 Observation – Quality of Implementation

Examining the effectiveness of TestEdge and the impact of the HeartMath tools cannot be adequately tested unless they are effectively implemented. The Time 2 observations and teacher interviews proved helpful in evaluating the implementation of the TestEdge program. Practice of the HeartMath tools led to a number of important successes, as documented by the quantitative and observational data, as well as in reports from the teachers. There is good reason to believe that sincere practice of HeartMath techniques had a substantial beneficial influence on physiological, psychological, academic and social aspects of the students' lives. However, the observations of the TestEdge program revealed that the program was not effectively implemented in its entirety, suggesting that the true potential of the program was not fully assessed.

The primary reason for the lack of a more effective implementation appears to stem from the fact that a number of teachers resented being required to add an additional program to their teaching load. There is little doubt that this affected these teachers' motivation and engagement with the program. The decision to implement the program was made by school administrators who sincerely wanted to expose students to the HeartMath tools. Since there was only a short time to complete the study due to grant requirements and the need to include the entire 10th grade population, school administrators determined that the English Language Arts class was the only option for providing the program to all the students at this grade level. Therefore, the 10th grade English Language Arts teachers were informed that they had been chosen to implement the program in their classrooms. It is unfortunate that this mandated approach had to be implemented at the site. Under the circumstances, however, there was no alternative.

A consistent theme of the observations was that many of the teachers were biased against the program before they even knew what it was; some actively resisted, and some simply were resigned to participating, as the following comments from teachers illustrate:

> "I don't think our staff was totally sold on it. It affected the attitude on it. We tried to do our best. Our kids, overall, weren't open-minded about it. … We have a very non-academic community. It's difficult to try to get out of the box."

> "This program came from on high so there was no question of administrative support. I had to field certain level of negativity from the teachers. … I may have been perceived as a "HM advocate" and perhaps they didn't open up with me."

> "It was really not fair to us or you—the circumstances we were in. Too abrupt! Changing an entire atmosphere (the school). (There was) poor communication between administration and teachers. This slowed down some people from getting passionately involved. A lot of people took issue with this project. Not me. Basically, this program is a band aid—a cover up for something which is fundamentally wrong. (We) need a deeper look at the issues."

Such sentiments kept some teachers from practicing the tools themselves. Most teachers felt that the TestEdge program took away valuable time they had already allocated to core subject material which they felt was essential and that this might reflect negatively on them, especially if their students did not do well on mandated tests. Many of the teachers were concerned about the curriculum they had to put aside. The following teacher comments are typical:

> "We had to slash and trim to make room for TestEdge. This was really not fair to us or to you."
>
> "I had to cut out part of the curriculum. I had a plan to stay on task. What was not covered in the understanding of the program was how much time it took to walk the kids back and forth to the labs. I made cuts but I had to make more cuts."
>
> "The way the administration gave it to us, without our 'buy-in,' impacted our attitudes. People resisted. Even when we were given the training, we were hoping it would go away."

Nonetheless, some teachers saw the need for such a program, were supportive, and described specific results, as the following comments from a teacher illustrate:

> "This program helped enhance students' intrapersonal skills—as they learned more self-knowledge. … This is a powerful way that we're dealing with emotional beings, and our emotions have a place here in the pursuit of success.
>
> [My students are] less anxious, more playful. [They are] more relaxed in class more often. Great camaraderie this year—there was an increase in it. The warmth they share, but maybe it's me, I'm different. … [They are] less apt to raise the ante from petty bickering to real fighting. Better anger management. … [They] exercise more patience for themselves and greater willingness to get the job done and less hesitancy to ask for that time. [They] made more of a connection about how they feel and how they perform. … My classes as a whole have shown tremendous calm during measurements of [test] performance. It doesn't mean they're better prepared. Every class has a 'squirrel' or two—class clowns that interfere. But overall, they've been more patient, and have more warmth.
>
> A few students have definitely gotten better …. If the [test] scores are up, then HeartMath can share in what's taken place."

Another important issue is the degree to which the tools taught in the program were effectively integrated into the daily routine of classroom life. Optimal outcomes from a program such as TestEdge derive from the sincere, consistent engagement of the tools by teacher and students, when integrated into the natural flow of classroom interactions and activities. The importance of such an approach was emphasized and modeled to the teachers during the teacher training sessions for the study. By the observers' accounts, however, there is good evidence that the tools were not consistently integrated in this way in the classrooms. Instead of modeling the tools for their students and integrating them into the entire classroom and school experience, most teachers only had their students use the tools as required by specific exercises in the program guidebook. For example, the following comments emerged during a debriefing session of the CGU observation team when reflecting on the patterns of TestEdge tool use they observed in the classrooms:[29]

> "The tools were not being practiced as a valued part of the class in May. If there were true integration, they would have been practicing in May. All the [TestEdge] posters were down. It was not part of the agenda. What they did as a part of each day was 'test prep'. I only saw *TestEdge* as a part of test-prep....They were very much following the booklet. They were only practicing the tools when they came up in the booklet."
>
> "… not once did I see a teacher, or someone suggest to a teacher, 'before you're teaching to this test' – which they all did – 'before you practice for this test, practice the tools; role model for them.' When you're practicing 'OK, we're going to do an essay today because these are the types of essays that are going to be on the test. But before we do, let's practice a Freeze-Frame; let's practice a Heart Lock-In; let's practice going to neutral.'… Each time…Then they'll get the experience of it. *And not once did I see this!*" (emphasis in original).

The effectiveness of teaching the TestEdge program and effective integration of tool use within the routine of normal classroom activities depends entirely upon the teacher. In the intervention school, nine teachers were responsible for implementation of the program. Drawing again from the CGU team's debriefing notes, it is evident from their observational experience that implementation of the program by most of the teachers was less than adequate and far from optimal:

> "I observed four teachers, who were excellent teachers and wanted to make sure the science of [TestEdge] was correct....But I don't think any of them practiced it daily, or used it in their own lives. I don't think any of them role-modeled the 'in the moment' type of teaching where if everything started going crazy they would have the students 'stop, and

[29] This section draws on Arguelles et al., pages 19-23.

> go to neutral.' I never witnessed it …. On the other hand … I think there were students who were doing it, that felt 'this is exactly what I've been looking for.'
>
> When I came back I expected to see them teaching English, teaching Othello or Caesar with the HeartMath tools intertwined – and I didn't. It didn't weave into their lessons. It still stayed a completely separate subject. I expected to see 'let's practice… now let's do the lecture… now, let's practice…now let's do some writing.' I didn't see any of that. It was all completely separate until the end. And it was the same in all the classrooms."

Schroeder observed one teacher who integrated using the TestEdge tools naturally into her teaching, while she saw three others who were less than effective:

> "As she taught it, she would do it. She would have personal stories of the things she thought were important about it, ways she had used it since she learned it. [This teacher] was serious about it. She asked for cooperation instead of punishing. She was doing a lot of heartfelt stuff. The focal point for her absolutely was the breathing. She was the only one I felt a sincere embracing and engaging of the tools."
>
> "…On the other hand, I think [a second teacher] hoped it would work, and wanted it to. But she was busy, and it was her first year, and she felt pressured to get what she had to get done. I think that's why she put TestEdge last. I think she thought it was a useful tool, but relegated it to last because she worried about getting everything else done. She was consumed. She didn't have tenure. She is a brand-new teacher and it became something else that was just added in. It was very stressful for her. TestEdge became part of the agenda; everything they *had* to do with the kids" (L.S.' emphasis).
>
> "[A third teacher] did the tools when they came up in the book and that was it."
>
> "[The fourth teacher] … no book, no breathing. She was a complete loss. I believe you can pretty much write it off. She was sarcastic about it. She didn't even do it with them; she fed it to them with a smirk on her face."

Overall, it appears likely that the impact of the intervention in this study was diluted by the lack of a sincere effort to teach the program by some teachers, inconsistency in integrating the tools into the daily routine of classroom life, the less-than-ideal timing and implementation issues mentioned above.

Many teachers suggested ways to improve the implementation, including starting earlier in the school year, having additional lab personnel, holding weekly HeartMath internet sessions, having students use the tools on other tests and writing assignments, and giving teachers time in the summer to integrate TestEdge into their curriculum. Other suggestions recommended placing the program in the career center, teaching it in a study skills course, or find a way of getting everyone committed without mandating the program. The following suggestions came from the Vice-Principal:

> Suggestions for Resilient Educator: "Divide into small groups of about three with one IHM person per group. Ask the (difficult) questions and discuss solutions. Then at the next training, restate the statements at the beginning of the training. Have the teachers be part of a weekly HeartMath internet group and do intellectual exercises. Create a real HeartMath network running weekly connect sessions, rather than waiting for weeks between contact. Have teachers embed HeartMath into every part of life."
>
> Suggestions for TestEdge: "Start in Spring of the previous year and identify one person to be the project leader and to train others. Identify the teachers and classes who will participate in project. Get teachers trained early or at least get them interested … and have them do exercises together so they would anticipate next year's implementation over the summer. Make it fun. Or, get them initially trained in August for the next year. Take time in summer to embed the TestEdge into their curriculum from the beginning of the year."

Overall, it appears that an adaptive approach (bottom-up) as opposed to a fidelity approach (top-down) to implementation may be more appropriate for a program like TestEdge (Arguelles et al., 2006). This perspective is supported by other studies which have found that teachers' emotional response to change is much more positive when that change is self-initiated, even when the change itself is mandated (Hargreaves, 2004; Dworkin, Larson, et al., 2003). This approach however, requires a stable administration with a sustained long-term commitment, which, unfortunately, is difficult to find in today's education environment.

In sum, the qualitative results portray a picture of overloaded teachers mandated to fit a program into an already crowded curriculum. Yet despite these obstacles, the quantitative data show that the majority of students were able to use the TestEdge tools to reduce their test anxiety.

Section Three
Results from
Study of Secondary Sites

Never before has accountability through rigorous content standards and standardized testing occupied so much of our teachers' and students' valuable learning time. Indeed, learning has been reduced to programmatic curriculum pacing and testing cycles throughout the school year. Teachers must keep up with the class next door, the class down the street, and the class across the state if they ever are going to cover the entire curriculum before the big test in the spring.

… The stress caused by this urgency at times can be overwhelming. To survive, teachers must limit their teaching to those subjects and topics that will be tested, namely reading and mathematics, while the other curricular areas fall by the wayside. The joy and deep satisfaction in sharing their love for learning with their students is usurped by mandates that are more appropriately designed for automatons and widgets. "There's no time to really enjoy great literature, to discuss author craft, or to ponder the significance of a book's illustrations. As soon as we finish one story we have to move onto the next if we are going to finish the theme in time for the unit test," lamented a third grade teacher.

—Jefferey Lagozzino (2006: 2)

Chapter XIV
Secondary Sites—A Case Study Approach

The secondary study, consisting of a series of separate qualitative case study investigations of schools in eight states (California, Florida, Delaware, Ohio, Maryland, Texas, Wisconsin, and Pennsylvania), was undertaken to evaluate the accessibility, receptivity, coordination, and administration of the TestEdge program across elementary, middle, and high schools and in school systems with different sociocultural, administrative, and situational characteristics.

Case Study Design and Methodology

The secondary sites and the grades in which the TestEdge program were implemented included:[30]

1) "Delaware Charter," a low socioeconomic status (SES), urban school with an all-African-American student population (grades 3 through 6);
2) "Maryland Elementary," low-to-middle SES, suburban school with a predominantly African-American student population (grades 4 and 5);

[30] To preserve confidentiality, the schools have been given pseudonyms, with each school named after its respective state.

3) "Texas Elementary," a low SES, suburban school with a 95% Hispanic student population (grades 5 and 6);
4) "Wisconsin High," a middle SES, rural, school with a predominantly Caucasian student population (10th grade);
5) "Pennsylvania Middle," a middle SES, suburban school with a predominantly Caucasian student population (8th grade);
6) "Ohio Elementary 1," an upper-middle SES, suburban school with a predominantly Caucasian student population (4th grade);
7) "Ohio Elementary 2," a low-to-middle SES, urban school with a predominantly Caucasian student population (4th grade);
8) "Florida Academy," an upper-middle SES, suburban charter school with a predominantly Caucasian student population (grades 6, 7, and 8); and
9) "Southern California Elementary," a low SES, suburban school with a predominantly Hispanic/Latino student population (3rd grade).

This component of the study utilized school and classroom observations, interviews, and observations of micro-outcomes to assess the process and context processes of implementation, teachers' personal experience with the stress management tools taught in the Resilient Educator Program, teachers' experience in introducing the TestEdge program to students, and any intended or unintended impact of the TestEdge program. This approach allowed us to uncover important processes, factors and dynamics that would have been difficult to assess defect quantitative methods. Thus, this study, in combination with the primary study, enabled a broader and more informed evaluation of the implementation of the TestEdge program.

All secondary study schools were provided a Resilient Educator training for teachers and other interested staff, Freeze-Framer heart rhythm coherence feedback systems, and with age-appropriate versions of the TestEdge program for use in the same manner as described for the primary study.

In the majority of cases, the observation of schools and classrooms was conducted by advanced doctoral students in education from Claremont Graduate University under the direction of the Co-Principal Investigator, Lourdes Arguelles, who was responsible for observational data collection and analysis at all of the primary and secondary sites.

The secondary sites and their selected classrooms were observed at two time points, one prior to and the other following the implementation of the TestEdge program. In depth, semi-structured teacher interviews were also conducted after the Time 2 observations and the completion of the implementation. Due to the diverse types of schools and grade levels and the limited number of classroom observations at any given site, a descriptive statistical analysis of means, and pre-post difference scores for each variable on the Classroom Observation Forms was not possible for the secondary sites. Therefore, we conducted an analysis of the primary constructs in the semi-structured interviews, the results of which were compared to the school and classroom observations. The observations were also used to better understand the school climate and the context within which the TestEdge program was implemented. However, "Southern California Elementary School," because of its close proximity to CGU, and therefore the availability of high levels of administration and teacher support, was observed in greater depth and will discussed as a case study of a site that was able to effectively implement the program.

Qualitative Analysis and Results

In the analysis of the semi-structured interviews with 14 teachers who taught TestEdge at the secondary sites, eleven primary constructs were assessed for response frequency. The first section of the interview focused on changes over the previous several years in the level of student fear and anxiety over taking tests. It also asked if teachers felt they had adequate preparation to help students manage such feelings. Only four (29%) of the fourteen teachers felt that before exposure to the Resilient Educator and TestEdge programs they had the requisite skills to help students with test fear and anxiety.

The second section of the interview inquired about teachers' experience with the Resilient Educator and TestEdge trainings, and asked them to rate the program on a scale of 0 to 10, with 10 being the highest. The mean score was 9.41, indicating that they found the programs highly relevant and valuable. There were no negative comments from any of the participants. The following are examples of typical comments:

> "The coverage of the material was excellent. The instructor showed us how it all worked. He modeled the tools, and that gave me confidence."

> "The training was really effective. I really enjoyed it and wish our whole staff had been there."
>
> "It helped me see how others, including students, can sense and respond to feelings."

The third section of the interview dealt with teachers' experience in implementing the TestEdge program. All of the teachers reported that it was easy for them to teach the program. The following are typical responses:

> "I found it easy to teach because the materials were designed at the level of the students."
>
> "When the kids saw me using the Freeze-Framer, they became interested and wanted to do it."
>
> "It wasn't very difficult to introduce the program. It was easy to follow and the kids loved working in the workbooks."
>
> "I found it very easy to discuss the information with the children because they could relate to the feelings or situations introduced in the program."

The next question probed the teachers' feelings about the time demands of the program. Only six (43%) of the fourteen teachers felt they had adequate time to teach TestEdge. Typical comments were:

> "We did the classwork once a week while I would have preferred to have done it daily."
>
> "There was not sufficient time for all the lessons. I had to rush through all the lessons to finish on time."
>
> "There was sufficient time, but time management had to be used creatively."
>
> "There was not sufficient time due to the fact that sometimes the scheduled advisory period in which the program was taught would get cancelled."

The last question in this section inquired about the level of administrative support for the program. Ten (71%) of the teachers indicated that there was support for the program; however, this did not always match the observers' perceptions. From their perspective, there was exceptional or adequate support in only four of the schools, and some support in one. In three of the schools the program was not well supported at all.

The fourth section of the interview addressed teachers' perceptions of the impact of TestEdge on student test anxiety, behavior, and academic performance. The first question asked if teachers had seen positive changes in student attitudes and behaviors. Nine (64%) of the teachers reported seeing improvements, while three reported they had not. The following are typical responses:

> "One of the major problems students experience is the feeling of being left out. Much of that changed during the TestEdge program. Students whose basic needs in life were not being adequately met began to open up more. When students used the TestEdge techniques to address these issues, they were able to generate calmer feelings."

> "There's lots of anger among the kids. They were very responsive to the Neutral tool. They would use it in different situations. I observed that the number of fights went down. They tended to think more about what they were doing instead of just reacting."

> "They loosened up more. They were especially excited about the Freeze-Framer. It has been a long time since I've seen kids this excited about something. It is a really unique tool. They felt it helped them with stress."

The interview also asked about teachers' perceptions of any change in students' ability to manage anxiety before and during a test. Seven (50%) said that they had noticed an improvement in at least some of their students, two reported that they had not observed a change, and the remaining teachers indicated that they could not answer the question because they had not yet had the opportunity to observe the class in a testing situation. The following are responses from teachers who noticed an improvement:

> "The students were more focused. We practiced the techniques before tests. Then, the students managed their stress during the tests by themselves. Students were more focused and calm."

> "The use of the breathing and the Freeze-Framer made a difference. If the students became anxious, I would let them go to the Freeze-Framer. Afterwards, they would go back and take the test whereas previously they would usually just give up or tear up the test."

> "I could tell the kids were using it. I saw some kids put their pencils down briefly and do the technique (breathing and generating a positive feeling) and then go back to the test. One test was particularly stressful because just before the test the electricity went out. The kids were talking and distracted and I started getting stressed. Then we all used the technique."

The next question in this section asked if teachers felt the program had impacted or improved students' academic performance. Eight (57%) of the teachers indicated that they felt the program had improved academic performance for at least some students while several felt that it had helped the majority of students. Three (21%) of the teachers indicated that they had not observed any improvements and one was unsure. Again, the following are typical responses:

> "I noted changes in a few students but I can't say they were all directly related to the TestEdge course."

> "Their FCAT (state proficiency test) performances were wonderful and I definitely saw them use the technique during the test."

> "The children tended to feel more secure or self assured and that tends to lead to higher scores."

> "I know of at least two boys whose performance went up."

The last question in this section asked whether teachers felt that the skills taught in the TestEdge program would impact their students' future development and academic performance. Ten (71%) of the teachers were convinced that these skills would have a positive impact on students' future development and behaviors. Four (29%) of the teachers said they did not know if there would be an impact. The following give an indication of the kinds of responses from teachers:

> "Many of the students have become pleased with making right choices and calming down when they need to. This will impact their lives. It has helped some already."

> "The exercises will help them with future test-taking."

> "I think that if a student practices the techniques and truly internalizes them they should improve in their future academic performance."

> "I believe that they now have tools to express themselves and manage their feelings. I think that it would be even more effective if they have teachers who could remind them of the strategies."

The last section of the semi-structured interview inquired about teachers' perceptions of the program's strengths and weakness, suggestions for improvement, and use of the tools and techniques in both their professional and personal lives. All respondents indicated that the program had helped them personally:

> "It has helped me stay calm during classes. The breathing exercise is helpful. I have also found that it helps me in traffic."

> "I am on blood pressure medicine. I come in early and use the Freeze-Framer. The games help me reduce my stress before class begins."

> "I use the techniques both in the classroom, and in my administrative responsibilities. I also use it in my personal life, especially so now that I am pregnant."

> "I have tried using the tools personally and feel that when I get myself to focus I can handle many situations better. I can't say that I have internalized them so much that they have become a natural part of my daily life, but I think the tools are very helpful."

> "For me as a teacher, the information about anxiety and effects of anxiety is very useful. Being able to express that in a way that is scientifically based rather than just me saying it is valuable as well. Having the tool to help the student get past some anxiety was just excellent."

The last question asked teachers if they planned to teach the TestEdge program in the following year's classes. All but one (93%) indicated that they planned on teaching the program to their upcoming classes.

An In-Depth Case Study: "Southern California Elementary School"[31]

In this section we report the results of an in-depth observational study of the highly successful implementation of the TestEdge program in a small K-3 school of economically disadvantaged and primarily English learner students in Southern California.

[31] Background Information and Observations on the Implementation of the TestEdge Curriculum in Third Grade at [Southern California] Elementary, Jefferey S. Lagozzino (2006), unpublished Qualifying Examination Paper, Claremont Graduate University.

The case study is noteworthy because it documents an exemplary implementation of the TestEdge program in challenging circumstances. It also shows that with proper preparation and under the right conditions, the program can be implemented successfully, producing strong, positive results in a relatively brief period of time.

The Research Site and Methods

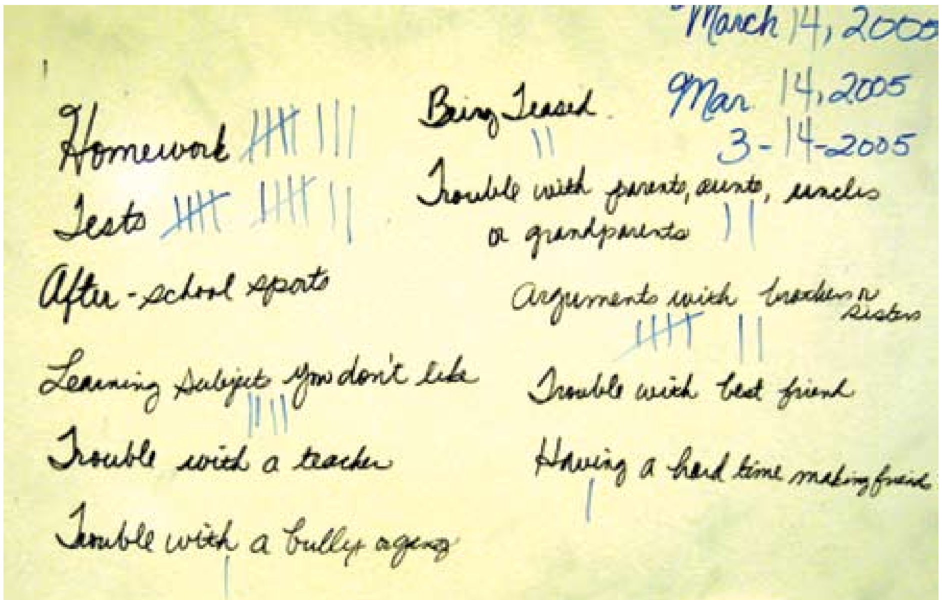
"Southern California Elementary School" is one of twelve schools in a district in Southern California. It is a small K–3 school with 14 teachers and 265 students. The student population consists primarily of English Learners whose home language is Spanish, like all schools in the district. All students are classified as economically disadvantaged, based on their qualification for the Federal Free/Reduced Breakfast and Lunch program. The ethnic composition is 83% Hispanic/Latino, 9% Vietnamese, 6% Chinese, and 2% other. Four third-grade classrooms with a combined total of four teachers and 61 students participated in the study.

The TestEdge program was implemented at the school in the Winter-Spring term of 2005. Because the TestEdge implementation at this site was the subject of an in-depth case study by Jefferey Lagozzino, a CGU doctoral student member of the TENDS research team, it was observed more often and in greater depth than the other secondary study sites. Observation and informal interviews were the primary methods used. The research benefited from the excellent principal and teacher support for the program. Observer records paint the picture of the challenges facing this elementary school and the benefits gained when the TestEdge program is effectively implemented.

"Tests Do Really Stress Them Out"

Data from observations and interviews conducted before the TestEdge intervention paint a picture of students *as young as seven* worrying about taking tests. In the course of a class discussion on the sources of stress in the students' lives, a third grade teacher conducted an informal survey and tallied the results on the board at the front of the class. The results (a digital photograph of the board) are shown in Figure XIV.1. The students identified eleven different sources of stress, most of which were school-related factors: homework, tests, after-school sports, learning subjects you don't like, trouble with a teacher, trouble with a bully or gang. The other factors were extra-curricular: being teased; trouble with parents, aunts, uncles, or grandparents; arguments with brothers or sisters; trouble with (a) best friend; and having a hard time making friends. Of the total of 37 tally marks shown on the board, 25 (67.6%) involved school-related stressors, the overwhelming majority of which involved either homework (8 marks, 32.0%) or tests (12 marks, 48.0%). *Overall, tests emerged as the most frequent source of stress (32.4%) in the young lives of these third-graders.* When reporting the results of her informal survey to Lagozzino (2006: 3) the teacher said, "I was really surprised by the results. Tests really do stress them out."

Figure XIV.1. A Third Grade Teacher's Informal Class Survey on Stress, March 14, 2005



In his informal interviews with parents, Lagozzino (2006) finds much concern over the emotional and behavioral consequences of their children's responses to the relentless pressure of state-mandated testing:

> "Are the kindergartners going to have testing soon?" a mother asked. "I know my niece in second grade is worried about taking the state tests. She's been talking about them since last week."
>
> When talking to Lagozzino, another mother says: "I didn't want (my children) to hear me talking to the teacher. My daughter has been getting headaches the last couple of weeks and it really worried me. I took her to the doctor and they checked everything out. They said she was perfectly healthy and thought maybe she was worried about something going on in her life. I was talking to her teacher about this.... She believes my daughter and three other children in the class are really stressed out about the state tests."

Lagozzino (2006: 3-4) also reports other incidents which show how extremely upsetting the testing regime is for some of these young children. Two are worth describing here for the uncontrollable physiological responses that such extreme emotional distress can cause:

> A second grade student, [Johnny], went into a tantrum outside his classroom door. His father was observably embarrassed by his son's surprising outburst. The child's emotion was so eruptive that it caused him to get a nosebleed. When the Principal arrived at the scene and calmly asked what was going on the youngster replied that he hated school and he hated tests. "I just want to go home."
>
> Another second grader expressed his disdain for the big test in a much different way. "You're not going to believe what happened," said his teacher. "[Miguel] was so upset about taking the test that he threw up all over it! What am I supposed to do?" The Principal had to call the district test coordinator who in turn called the state department of education. The official directions were that the soiled test needed to be sealed in a zip-lock bag and returned with the other tests due to test security measures. The child was issued another test booklet to take the test and luckily was able to control his urge to vomit throughout the remainder of the testing period.

In commenting on his observations, Lagozzino (2006: 4) is careful to note that "while not all students respond in the same way or with the same intensity," what these "scenarios" clearly show is "the intense emotional and physical harm test stress can cause." What is most disturbing about the high-stakes, high-pressure testing that No Child Left Behind has mandated is the emotional and physical harm it has wrought on our youngest and most emotionally defenseless students. In this situation in which there is constant pressure on schools and students to achieve ever-increasing test scores, "it behooves educators to equip their students with the emotional understanding and tools to deal with this stress in productive, healthy ways."

TestEdge Training

As the first step in the implementation of the TestEdge program, each teacher attended three two-hour workshops in February 2005. The workshops were held after school and presented by an Institute of HeartMath trained facilitator. The goal of the workshop was to instruct the teachers on the science behind the TestEdge program and to teach them the program's tools and techniques so that they could practice and apply the stress-reducing strategies in their own lives. This would provide them with the requisite knowledge and practice experience to teach these skills to their students. In addition to participating in the workshops, to facilitate their practice of the tools and techniques, all teachers were provided with a copy of the Freeze-Framer technology to install on their home computers.

In the first teacher workshop, there was a general "air of nervousness and discomfort" among the participants, who "were hesitant to respond" and share their own personal experiences, despite how well they knew one another. However, during the second workshop, which involved learning how to use the Freeze-Framer technology, there was a feeling of great interest and shared enthusiasm among the participants (Lagozzino, 2006: 6).

> Once the computers were turned on and the teachers connected the heart rate variability monitors to their fingers they were "hooked" by the technology. The mood in the room was lively and energetic.… [T]he participants could not contain their interest in seeing how each other was doing which initiated a spirit of good-humored competition amongst them. "Look at [Kathy's] screen," the Principal whispered to the School Psychologist. "She's been 100% in the green (indicating a state of [high] coherence) since she hooked up!" "That doesn't surprise me, does it surprise you?" she asked. "That's just [Kathy]. It's her personality. She's always so calm and collected. Nothing really seems to agitate her. I've sensed that from her from the first time I met her." "No, it doesn't surprise me either. It just confirms how I see her too," replied the Principal.

After two weeks of practicing the new tools and techniques on their own, the four teachers were given the final workshop, which focused on instructing the teachers how to teach the TestEdge program to the students in their classes. The teachers were familiarized with an elementary school version of the TestEdge curriculum designed for grades 3-6. This curriculum consists of fourteen 30-minute lessons supported by the following materials: poster-sized lesson cards, a student workbook, and an optional Freeze-Framer system. To implement the program, the teachers agreed to teach one or two lessons a week to cover all of fourteen core lessons before the spring testing was to occur in May. They felt that they could "easily" accomplish this within the twelve weeks available.

TestEdge Implementation

Overall, the implementation of the TestEdge program by all four teachers was successful, as the following observations and comments from interviews show (Lagozzino, 2006: 7):

> "The students love the workbook. It's really well done and directly reinforces what they are learning in the class lessons. I've been so amazed at how they are applying what they are learning to other situations. The other day I was reading a story aloud and some students commented that the character should have done a Freeze-Frame[32] to help her with her problem," commented one teacher.
>
> "Once the technology was introduced to the kids, they were hooked. It was a little slow going in the beginning until we got to the Freeze-Framer. Now when we're reading a book aloud the kids will comment about a character and say, 'she should do heart breathing and do a Freeze-Frame.' They're starting to see situations in which the tools can be helpful," remarked a third grade teacher.

In his observations around the school and informal chats with students, the Principal also found that there was a positive response among students to the TestEdge implementation, especially when they were using the Freeze-Framer technology (Lagozzino, 2006: 8):

> "Have you guys been talking about stress in your class?" asked the Principal. "Yeah, we're learning how to do better on tests," said a third grade boy. "Have you used the Freeze-Framer?" asked the Principal. "Yeah, it's cool. The first time I was in the red [low coherence] the whole time. It was hard. The second time I got a lot of blue [medium coherence]," stated the student. "How did you do that?" asked the Principal. "Oh, I just breathed and was thinking about being happy and it was blue. It was really cool."
>
> On another occasion the Principal asked a group of four girls who were participating in the study about what they were learning in class. "What are you learning in class about taking tests?" the Principal asked. "We're learning how to breathe and relax," answered one of the girls. "Have you used the computer to help you?" the Principal inquired. "Yeah," all four girls replied in unison. "I was only able to get the waterfall in color when I tried,"[33] commented one. "The teacher was able to get everything

[32] Freeze-Frame is one of the techniques taught in the TestEdge program to aid in problem-solving.
[33] Referring to the Freeze-Framer's "Meadow Game," in which a black-and-white meadow scene fills in with color as the user increases and maintains coherence.

> in color!" remarked another excitedly and then continued, "Our teacher said that we need to practice heart breathing at home and I've been practicing. I hope I'll be able to get more of the color picture next time."
>
> On one visitation to a participating classroom when a substitute was teaching, the Principal observed that many students chose to use the Freeze-Framer on the computer as a free choice activity after completing their regular class work.

The School Psychologist reported having similar success with her five-year-old son in teaching him how to use the Freeze-Framer technology at home (Lagozzino, 2006: 8-9):

> "He'd tell me, 'Okay, now I'm thinking about playing with my friends. It's fun!' and, 'Now I'm thinking about this thing or that thing.' He was doing really well. Then my daughter ran into the den and screeched 'mom' and his coherence level went immediately from high to low. At which point my son yelled 'Emma! You messed up my happy place!' It was really cute and very interesting. He couldn't get back to high coherence with his sister in the room."

An important strength of the TestEdge program that teachers commented on was the relative ease with which it could be incorporated into the existing curriculum. Lagozzino reported that the teachers found that TestEdge provided "opportunities to move beyond simple test preparation techniques and could be fully and naturally integrated with other areas of the curriculum." Given the enormous pressure of the existing curricular demands on teachers, it is important when implementing a new program such as TestEdge, as Lagozzino notes (2006: 7), that it can be readily incorporated into the existing curriculum, rather than having to be taught as a separate, discrete topic.

Another strength that teachers noted was the integration of the optional technology component. While this proved to be "both challenging and rewarding," all of the participating teachers agreed that it was integral to successfully teaching the program, "even to students of this young age." One teacher felt that the computerized Freeze-Framer technology was the "key to piquing" students' curiosity about feelings and emotions and sparking their interest on the topic of emotional management (Lagozzino, 2006: 7).

Outcomes of the Intervention

On May 9, 2005, the students in second and third grades underwent nine days of standardized testing mandated by the state. Lagozzino examined his anecdotal qualitative data and student attendance records as a way of gauging the effects of the TestEdge intervention on the third grade students and their teachers. Although it would not suffice as a scientific control, he used the outcomes for the second grade students and their teachers as an informal baseline for comparison.

Starting with the results on student attendance during the 9-day testing period, Lagozzino (2006: 10) found that there were a total of 29 tardy arrivals and 8 absences for the 49 students in the second grade, compared to 6 tardies and 4 absences for the 61 students in the third grade. Aggregating tardies and absences together and dividing the product by the total attendance days for each grade, there was an 8.4% deficit in perfect attendance for second grade students (control) compared to only a 1.8% attendance deficit for the third grade students (intervention). While not definitive, these results are consistent with the expected effects of the intervention.

In addition to the difference in student attendance during the testing period, there was also evidence of differences between second and third grade teachers in terms of stress levels, attitudes, and observations of their students' behavior, as Lagozzino (2006: 9-10) reports:

> Each day the teachers had to come to the Principal's office to sign out their testing materials and return them each day to the same location when they completed that day's test. The third grade teachers' attitudes were calmer than their second grade colleagues. Their demeanor was not noticeably different than any other regular teaching day. However, the two second grade teachers seemed much more agitated than they are usually. "Why do we have to check these tests in and out everyday? Do they think we're going to cheat and tell the kids what's on the next day's tests? This is ridiculous!" curtly stated one second grade teacher. These comments were surprising in light of the fact that the same procedures have been in place for several years now.
>
> When the other second grade teacher was asked about how his students were doing with the test, he exasperated, "All you can do is encourage them to do their best. It's very frustrating to see how they work quickly just to get the test done. You ask them to double check their answers and they just sit there until you collect the test from them."

> In contrast, third grade teachers had much different comments on their students' approach and performance on the tests. "They seem very relaxed. I remind them of the tools that they can use before we start each day. Today, the students were even joking around pretending they were stressed out before we started and laughed it off like taking the test was no big deal," one teacher related.
>
> "They are really taking their time to do their best. I'm really impressed with their perseverance. Since today's test was un-timed more than half the class was still working when the recess bell rang so I told them to continue working and we'd have a later recess on our own. Normally when the bell rings they're ready to go. Their minds have switched off, but many of them sat for another thirty minutes or so to finish going over their answers. I saw several students re-reading passages and checking to see if they could match their answer with information they had read. The others just sat quietly until the last students finished," commented another teacher.
>
> Taking extra time when it was available was observed by all of the third grade teachers. While not every student took advantage of this opportunity, the teachers felt that a greater number of students did do so than in the past. "I think that the students who took the extra time were really focused on doing their best. I'm really proud of them," shared one of the teachers.

One of the third grade teachers reported observing greater emotional empathy and support among his students during a debriefing session after a testing day:

> "They have been very observant of not only their own feelings but also of their peers' feelings ... . [T]he students told one of my girls that she should do a Freeze-Frame when she's testing. He told her that he noticed how nervous she was because of the way she was putting her fingernails to her mouth and demonstrated this to her. The other students told her that they noticed her doing this too. She took the advice very well. Testing has been very difficult for her even though she is so bright," commented the teacher. "I thought it was very caring the way the class was so empathetic to her feelings about testing and trying to give her emotional support."

Finally, in response to the question, "Would you use the TestEdge curriculum again next year?" all of the participating third grade teachers responded enthusiastically in the affirmative. One teacher replied:

> "Definitely! It teaches about emotions in such a meaningful way. It's so much better in discussing emotions than our mandated curriculum

> on drugs and violence. I would start earlier in the year. Right from the beginning, so that the students could have a lot of time practicing the tools before testing."

Test Results

After the test results were released, a HeartMath researcher interviewed the school's Principal to find out how the third grade students had fared. He stated:

> "Our test scores soared, far exceeding the NCLB targets for the school year. In Language Arts the students' proficiency grew from 26% to 47%, and in Mathematics from 60% to 71%."

He went on to say that he is convinced that the TestEdge program was "*the* significant factor" in increasing his students' scores on the state-mandated tests.

The Principal also said that he is eager to continue the program and make HeartMath programs available to other members of the teaching staff and to parents. He is convinced that effective emotional management training for children should begin early, and that it begins with exemplary adult role models.

When told of the Principal's desire to have them teach the TestEdge program again next year, all of the participating teachers affirmed that they would like to and said that they felt that the curriculum and tools were valuable for their students to learn as well as an effective way to manage their own stress levels.

Conclusion

This case study demonstrates that with adequate preparation and under the right conditions, the TestEdge program can be implemented successfully and produce strong, positive results in a relatively brief period of time. Needless to say, finding a sympathetic and cooperative administrative and teaching staff is critical. When there is unconditional endorsement of the program from key school administrators, when sufficient time and care are taken to introduce the program to the teaching staff, and when appropriate support is provided at all levels, students not only respond enthusiastically to the new concepts, but they quickly integrate the HeartMath tools into their school and personal lives. The result is improved academic performance as well as improved attitudes and behaviors.

Summary and Discussion of Findings

There were a number of important findings from the series of secondary site case studies. To begin with, there was a universal recognition by the teachers that the lack of emotional self-management education for students resulted in significant blocks to learning. They felt that the time necessary to provide such education would be well worth the investment since it would save tremendous amounts of time by reducing classroom disruptions, improving student focus, and helping students learn and remember new content. Very few teachers felt their previous training had given them the skills to deal with stress, let alone help their students do so.

All of the teachers at the secondary sites reported that they personally benefited from the Resilient Educator program and most felt they would benefit from additional training or follow-up sessions. Many of the teachers and school administrators expressed the wish that the program be made available to all school staff. Based on the teachers' comments and the observers' impressions, providing teachers with the Resilient Educator program well in advance of their introducing the TestEdge program to students would make for a more effective implementation strategy.

In general, both teachers and administrators felt that implementation of the TestEdge program in this study was too abrupt for a smooth integration into the classrooms. This was due in part to the fact that grant approval was obtained after the school year had already begun and that grant guidelines encourage completion of the study in a one-year period. In general, the implementation and integration of the tools into the classroom was more successful when there were several teachers at the same grade level teaching the course. At these sites, teachers had a built-in support system where they could discuss approaches and applications with one another.

A common observation was that some teachers felt they were not able to adequately help students make connections between the tools they were learning and specific situations in which they could be applied. It was apparent from the data that at sites where teachers were able to internalize the use of the tools in their own lives, they were better able to help students see how they could use the tools in their lives. These were also the sites where teachers reported improvements in student behaviors and performance and expected that the program would have long-lasting benefits for the students. Another conclusion is that teaching two TestEdge lessons per week produced better results than teaching a lesson every day or only one lesson per week. This seemed the best frequency to allow students to integrate the lesson material until the techniques become familiar and automatic.

It was notable that almost all teachers at the secondary sites commented on the value of the Freeze-Framer technology, both for themselves and for their students, especially in helping students internalize the use of the tools in such a way that they looked forward to the experience. This is consistent with the observations at the primary intervention site, where students chose to use the technology on their own time in the periods before and after school.

Another common perspective was that additional time needs to be allocated to teacher orientation and logistics. There were a number of logistical challenges in some of the schools. In some cases there was ineffective communication between administrators and teachers. Also, a surprising number of teachers were not computer-literate and therefore had difficulty using the Freeze-Framer system.

It is clear that additional implementation support in some schools would have led to more positive outcomes. The support of the principal and other key school administrators cannot be overstated. It fact, it was likely the most important factor as to whether or not the school had a successful implementation.

Section Four

Discussion
and
Conclusion

Chapter XV
Findings and Limitations

The large volume of quantitative and qualitative data collected from multiple sources has yielded a rich bounty of empirical results described in detail in Chapters V through XIV. Viewing the findings from a broader perspective, the primary concern of this chapter is to identify the study's major findings and to weigh their greater significance in relation to limitations in the study's design, implementation, methodology, and data analysis procedures.

In sifting through the many results to identify the major findings and to guard against the problem of spurious inference, we were guided by the principle of consistency of evidence which we applied in evaluating both the quantitative and qualitative data. This principle holds that the likelihood of valid inference—that a given empirical finding is true—is increased insofar as there is a consistent pattern of evidence from multiple sources, measurements, and/or analyses. The validity of inference is further enhanced to the degree that the finding is consistent with the *a priori* expectations of an explicit theory or hypothesis and/or is consistent with the evidence of prior research.

Before presenting our discussion of the major findings, we will begin with a brief overview of the study's research goals and design.

Research Objectives and Design

This study was initiated to determine the magnitude, correlates, and consequences of stress and test anxiety among elementary, middle and high school students and to investigate the efficacy of the TestEdge program in reducing students' stress and test anxiety and improving their emotional well-being, quality of relationships, and academic performance. In addition, we set out to characterize the receptivity, coordination, and administration of the TestEdge program in a wide variety of school systems with different cultural, administrative, and situational characteristics.

Two major hypotheses were tested in the study. We expected that—all other things being equal—(1) Competence in practicing the TestEdge tools would result in significant improvements in emotional self-regulation and psychophysiological coherence. These changes would produce a marked reduction in test anxiety, which in turn would lead to a corresponding improvement in academic and test performance; and (2) Given the logic of Hypothesis 1, we also expected that there would be concomitant improvements in emotional well-being and stress management, life aspirations, behaviors, and relationships in the student's life as a whole, as well as improvements in classroom climate, organization, and function.

To test these hypotheses we implemented two studies, each with different objectives and designs. The primary study utilized a quasi-experimental, longitudinal research design involving experimental and control groups, with pre- and post-intervention measurement within a multi-methods framework. It employed survey questionnaires, structured and semi-structured interviews, student drawings, classroom observations, and electrophysiological measures. The secondary study consisted of qualitative investigations of implementations of the program in schools in eight states (California, Delaware, Florida, Ohio, Maryland, Texas, Wisconsin, and Pennsylvania). These investigations employed a case study approach from which observational and interview data provided valuable information on best practices and pitfalls when implementing interventions such as TestEdge in schools with diverse compositions and characteristics.

Findings from the Primary Study

Due to the large volume of quantitative data collected in the primary study and the many hundreds of statistical analyses we conducted, we faced a basic question of statistical inference—being able to discern valid findings from results produced by chance. As discussed in Chapter IV (see pages 57–59), the risk was commission of Type 1 and Type 2 errors of inference: accepting a statistical result as true when, in fact, it is invalid; or rejecting a result as false when, actually, it is valid. As indicated at the outset of this chapter, we used the principle of consistency of evidence in evaluating statistical results in order to keep these errors of inference to a minimum. While this worked well when interpreting results beyond the level of chance, we were more vulnerable to spurious inference when dealing with results close to the threshold of chance (within a few hundredths of $p = 0.05$). Thus, as a further aid to our decision as to whether to accept or reject a given marginal result as valid, we followed three additional guidelines. Although presented earlier in Chapter IV, we will take a moment to briefly review these guidelines because of their importance in what follows.

Guidelines for Statistical Inference

To accept a result of marginal statistical significance as valid, the result must meet either of the following two requirements:

- The first is that the result must be consistent with prior research and the study's major theoretical expectations. This guideline rules out a marginal result that does not make sense within the study's theoretical interpretive framework.
- The second requirement is that the result must be consistent with an existent pattern of findings from the study or must be corroborated by data from another source or method. This is the criterion of internal consistency, which holds that there must be some pattern of evidentiary support for the result, beyond that of the result itself.

From our analysis of subgroups within the sample and our analysis of matched-case/group comparisons we have many results well beyond chance involving small case counts. What this means is that we are vulnerable to any sampling error involved in the multivariate selection procedures used to construct the comparison groups.

- To guard against errors of inference from this source, our third guideline was to treat such results with caution, regarding them as suggestive or preliminary rather than definitive.

Findings from the Whole Sample

In the analysis of all students who participated in the primary study (from both schools), we learned that, in general, students planned on graduating from high school and felt that there is a purpose to their life (Educational Plans), and they felt safe and happy in their respective schools (Feelings About School). They also perceived that they could talk to their parents if they had a problem and that their parents were supportive and interested in their schoolwork (Parental Support). As well, they felt fairly positive that what they were learning at school would help them succeed as adults and prepare them for college or getting a job (Life Preparedness). Students generally felt that they had a strong network of friends (Extent of Friendship). In general, they rated themselves as happy, enthusiastic, and hopeful (Positive Affect), and they felt that there was at least one teacher or adult at school who listened to and cared about them as individuals (Teacher Support).

On the other hand, they indicated that they were not able to manage stress or effectively manage feelings of anxiety or being overwhelmed (Stress Management). This was related to their tendency to repress negative feelings and feel powerless over their emotions (Emotional Discord). In addition, they felt it was difficult to know what other people are feeling, tended to feel misunderstood by others, and found it difficult to keep from getting into arguments or fights (Interactional Difficulty).

There was also evidence of a significant experience of test anxiety for a large portion of the sample. Of the 749 10th graders sampled in the study, 61% were affected by test anxiety, with 26% experiencing high levels often or most of the time. We also found a strong relationship between test anxiety levels and test performance on the two CAHSEE tests and on each of the California Standards Tests. On average, students with high levels of anxiety scored 15 points lower in both Math and English-Language Arts than the students with low test anxiety. Given the socio-economic difference in community context, and the differences in ethnicity and academic performance between students at the two schools, it was somewhat surprising to find that students at both schools had equally high levels of test anxiety.

These numbers are indicative of a serious problem and justify the concern that test anxiety may significantly jeopardize the assessment validity of tests and constitute a major source of test bias. To the extent that anxiety inhibits performance, the testing process does not measure a student's true competence or ability.

Consistent with previous research, we found that students with high academic ability (those in advanced classes) feel more supported at home and school, have lower levels of test anxiety, and are more positive about school and more optimistic about their futures than students of average ability. By contrast, average students worry more about taking tests, and they report higher levels of stress, loneliness, disappointment and depression (Negative Affect), as well as higher levels of emotional discord and interactional difficulties.

As have others, we also found that as a group females suffer more from test anxiety than males. If fact, we found that twice as many girls experience high levels of test anxiety compared to boys. Some investigators have theorized that this is because girls are enculturated to be aware of and express feelings, whereas boys are taught to deny and suppress them (Cizek & Burg, 2006). However, although this may be a factor in some cases, our finding is in line with numerous clinical studies which have found that young females are more than twice as likely to suffer from anxiety and panic than males. Female students in our study also reported high levels of stress, loneliness, disappointment and depression (Negative Affect) as well as emotional discord. On the other hand, they also reported higher levels of positive emotions, felt more support from teachers, had deeper close friendships, and were more optimistic about their futures than were the males.

In agreement with other studies, we also found that Hispanic students worry more about taking tests, feel less supported by a teacher or an adult at school, and experience less parental support than their White classmates.

In regard to the relationship between sources of student stress and academic performance, we found that five common factors in a regression analysis model explain approximately 20% of the variance in student test performance on both the CAHSEE and the CST: 1) the worry component of the Test Anxiety scale, 2) how close students felt to others and how safe and happy they felt at school (Feelings about School), 3) the emotionality component of the Test Anxiety scale, 4) students' perceptions of how much school would help them be successful in life and get a job or go to college (Life Preparedness), and 5) their feelings of having a purpose in life and intent to graduate from high school (Educational Plans). These findings offer some support for both of the study's hypotheses, and they are consistent with previous studies which have also related test anxiety and socioemotional factors to academic performance.

An independent analysis of the TENDS data (Hartnett-Edwards, 2006), using multiple regression analysis, found that measures of Affective Mood explained approximately twice the variance in student test performance on both the CST-ELA and

CAHSEE-ELA as items from the TAI (23% versus ~13%, respectively), and that measures of student Social Behavior have about the same predictive power (~12%) as measures from the test anxiety scale. It was found that positive feelings and emotions and prosocial behaviors have a positive effect on test performance, while strongly negative feelings and antisocial behavior and interactions have a negative impact. These findings suggest that students' overall emotional disposition—their emotional awareness and skill to appropriately manage emotions and feelings in their lives as a whole—appears to be a *stronger* predictor of test performance than their specific fears and worries about taking an important test. In addition, they suggest that the nature and quality of students' relationships, both in and outside the classroom, are *equally* important in affecting test performance as is anxiety about a particular test. These results are also consistent with the theoretical expectations outlined in Chapter II, namely, that student emotions and the classroom socioemotional environment both exert an important effect on student learning and test performance.

The Question of Intervention Effects

In moving to a discussion of the key question of intervention effects, we began the analysis by examining the equivalence of the intervention and control schools. Importantly and somewhat surprisingly, given the differences between the two schools, there were no significant differences between the students on any of the measures of test anxiety. However, four factors were identified for which there was a significant difference. The first was test performance, with students in the control school outperforming those in the experimental school by a notable margin. These students also had closer relationships, felt safer and happier (Feelings about School), and were more connected with their network of friends (Extent of Friendship). However, students in the experimental school felt that school would help them be successful and get a job or attend college to a greater extent than students in the control school (Life Preparedness). With the exception of these four factors, there were no significant differences at baseline between the two schools.

These differences in baseline test performance were not unexpected because of the control school's general more favorable history of academic achievement. However, they did have implications for the analysis strategy, requiring us to match subgroups of students from the two schools on 9th grade test scores to control for the differences.

An important finding by the end of the study was significant reductions in all three measures of test anxiety (Global, Worry, and Emotionality) in the experimental as compared to the control group. At the beginning of the study, 61% of the students at the intervention school reported being affected by test anxiety; by the end of the study, 75% of these students had reduced levels of test anxiety. Moreover, a pre- to post-intervention reduction in mean test anxiety was observed in 76% of the classes in the experimental group; by comparison, only 20% showed a significant decline in test anxiety in the control group. What is especially notable about this finding is that this pattern of test anxiety reduction in the experimental group classes occurs throughout the test performance spectrum, from low-scoring to high-scoring classes.

The experimental group also showed significant reductions in feelings of loneliness, sadness, anger, depression, and disappointment (Negative Affect scale) and significant improvements in their emotional awareness, power to control their feelings, and measures of Emotional Discord (i.e., they were not as overwhelmed by their emotions and were less inclined to repress negative feelings). Consistent with the expectations of the second hypothesis, they also had significant improvements in their awareness of others' feelings, their ability to share their feelings with others, their feelings of being understood, and their ability to avoid arguments or fights (Interactional Difficulty scale). In addition, there was a significant increase in students' enjoyment of and learning in class, positive feelings towards classmates, and increased perception of teacher care (Positive Classroom Experience scale). Finally, for the students in the experimental group there was a strong positive relationship between frequency of TestEdge tool practice and increased use of tools across different life situations.

The notable reduction in test anxiety and the improvement in emotional self-regulation and relationships are consistent with the expected effects of the TestEdge intervention and support several key elements in the two hypotheses tested in the study. Specifically, this was true of the first hypothesis' prediction that the tools taught in the program would be related to improvements in emotional self-regulation and reductions in test anxiety. The other key elements of this hypothesis state that emotional self-regulation would be related to increased psychophysiological coherence, and that the above changes would be related to greater improvements in student test performance. As we will discuss shortly, the prediction of increased coherence was clearly realized, and there is evidence that this is related to improvements in test performance, although the latter is not as strong and direct as we would have hoped.

Due partly to the differences in academic performance and socio-economic level favoring the control school over the intervention school (described in Chapter V), and to the complexity and non-linearity among the expected relationships between reduced stress/test anxiety and test performance, only a small degree of nonsignificant change was observed between the experimental and control schools' test performance for the entire 10th grade population as a whole. Therefore, our analysis turned to an investigation of subgroups within the experimental and control groups where a matched-groups approach was used to control for the baseline differences in test performance and likely associated sociodemographic factors.

We used two matched-group strategies. One involved the construction of matched groups on various combinations of sociodemographic variables and the SOS scales. The second matched classrooms of students on their 9th grade CST test performance (baseline). The matched-groups approach not only revealed that a number of subgroups in the experimental school evidenced a reduction in test anxiety, which was also associated with improvements on measures of emotional disposition, but also found that four subgroups had achieved significantly higher test scores.

The first group consisted of 82 White females in average academic level classes. Results from the pre–post-intervention analysis showed that these girls had both a significant reduction in test anxiety and significantly higher test score gains—on average by a margin of approximately 13 points—in their CST English-Language Arts tests than their classmates in the control school. This finding is noteworthy because it suggests that the TestEdge program is readily accessible and helpful to students within normal levels of academic ability.

The second group also consisted of females. However, this group of fifty girls had ethnic identities in the "Other" category (Asian, African American, Pacific Islanders, etc.). In addition, they were in an advanced level class. In this case, the girls in the experimental group outperformed their control group counterparts by a margin of almost 25 points. However, this improvement in test performance was not associated with a corresponding reduction in test anxiety.

For the third and fourth subgroups, classes of students were matched on their baseline (9th grade CST) Math scores as a way of approximately matching students on their intellectual ability. The first group, Math Group 1, contained 129 students. Like the White females in a regular academic class, this group also had a significant reduction in test anxiety, which was associated with a 10-point higher average test score on the experimental school students' 10th grade ELA tests.

The second group, Math Group 2, was comprised of 67 students. In this group, the experimental school students had a significantly higher performance on their 10th grade CST Math test than did the control group—by a margin of approximately 23 points. They also scored approximately 9 points higher on their 10th grade CST ELA test, although this difference did not quite reach statistical significance. This increase in test scores was not associated with a significant decrease in test anxiety.

In addition to these findings, we found a number of commonalities in emotional and relational disposition that appear associated with the changes in test anxiety and test performance. First, a reduction in test anxiety is likely to be associated with an increase in positive affect. Second, reductions in test anxiety are also usually associated with a reduction in negative affect, emotional discord, and interactional difficulty.

Findings from the Physiological Study

The physiological study was designed to provide an objective measure of the degree to which student ability to manage test stress had improved after they had taken the TestEdge program. The study was designed as a controlled laboratory experiment using continuous electrophysiological measurement of HRV during computer administration of the Stroop Test—a word-color conflict stress test. The effort to construct equivalent experimental and control samples, by randomly stratifying the selection of participants, was only partially successful due to conflict with student class schedules. This resulted in an over sampling of students from advanced classes in the control group.

There were a number of important findings from the physiological data that indicate that even though the tools in the TestEdge program were introduced to the students in a less than ideal way, students nevertheless appear to have practiced and used them on a regular basis to manage their stress in a variety of situations.

For the group as whole, the students who learned the coherence-building tools showed an overall HRV increase during the four-month period between the first resting baseline recording, measured in January (Time 1), and the second, measured in May (Time 2). It is noteworthy that the students in the experimental groups showed a significant lower overall HRV at Time 1 than the control students, a pattern that was reversed by the Time 2 recording. This is a clear indication that an improvement in autonomic nervous system function occurred in the students in the experimental group. This finding has important implications, since HRV is considered a psychophysiological marker of emotional self-regulation abilities and core regulatory functions. In addition, low HRV, especially the aspects reflecting parasympathetic activity, is associated with a loss of inhibitory control of anxiety. As far as we know, this is the first study to show that high school students' HRV can be increased over a relatively short period, let alone through a classroom-based program.

Not only did the experimental group students' overall HRV increase at the Time 2 baseline recording, but also their ratio of heart rhythm coherence significantly increased. This finding provides further evidence for the efficacy of the TestEdge program—in that using a systematic process to self-regulate emotional experience and shift to a state of increased heart rhythm coherence on a consistent basis effectively facilitates a repatterning process. In this process (described in greater depth in Chapter II), coherence becomes increasingly familiar to the brain and nervous system, and thus progressively becomes established in the neural architecture as a new, stable psychophysiological baseline or set point. Once the coherence mode is established as the new familiar pattern, the system then strives to maintain this mode automatically, thus rendering coherence a more readily accessible state during day-to-day activities, even in the midst of stressful or challenging situations.

Consistent with the physiological indicators of increased HRV and coherence, there was a significant reduction in feelings of loneliness, sadness, anger, depression and disappointment (Negative Affect scale) in the experimental group, which was accompanied by a large and significant reduction in all three scales of test anxiety.

There were also a number of significant findings from the stress preparation phase of the experiment. This was the main focus of the physiological study, since it was intended to measure the degree to which students had acquired the ability to intentionally shift into the psychophysiological coherence state prior to an upcoming challenge or important test. The main finding was that the experimental group's heart rhythm coherence ratio was substantially larger than that of the control group during this phase of the experiment, and that it was substantially higher than that observed during the resting baseline period. In addition, the experimental group had lower heart rates than the control group, and they had significantly greater high frequency power and low frequency power in the HRV power spectrum.

When combined with the findings from the resting baseline period, the experimental group's shift to higher coherence during the stress preparation phase is notable. It suggests that not only did a substantial number of students learn how to shift into coherence, but also that to be able to do this they *must* also have practiced the skills. HeartMath has found that for most individuals, it takes some degree of practice to learn how to make this shift at will and to be able to sustain the coherent mode. Thus, the stress preparation results indicate that students in the experimental group had learned and practiced the coherence-building skills—in many cases, even without the requisite encouragement and support from their teachers—either by using the Freeze-Framer system in the computer lab in class, by working with the program guidebooks, and/or by practicing the use of the skills in various life situations.

The power in the high frequency band of the HRV spectrum is widely accepted as a measure of parasympathetic or vagal activity, which is often associated with the relaxation response. The peak in the low frequency region can reflect both sympathetic and parasympathetic activity, especially in short-term recordings. However, parasympathetic influences are particularly present in this band when respiration rates fall below seven breaths per minute. In the coherent state, heart rhythms, blood pressure rhythms and respiratory rhythms synchronize at a rhythm around six cycles per minute. When this synchronization occurs, the HRV pattern becomes a sine wave-like signal and the HRV power spectrum shifts to a narrow-band, high-amplitude peak centered in the low frequency region. Because the frequency at which coherence occurs falls in the center of the low frequency region, it could be misinterpreted as an increase in sympathetic activity.

However, in the case of the coherent state, the distinctive shift to a narrow-band, high-amplitude peak centered in the low frequency region is primarily due to increased order in higher-level control systems in the brain, increased synchronization between the two branches of the ANS, and a general shift in autonomic balance towards increased parasympathetic activity (McCraty et al, 2006). This means that the coherent state encompasses the increase in parasympathetic activity associated with increased relaxation, and also reflects increased order and synchronization in the activity and functions occurring in the brain and nervous system. This increased order and harmony has been hypothesized by neuroscientists and psychophysiologists to be a causal factor which enhances learning and performance (McCraty et al., 2006; Ratey, 2001). For example, increased coherence is associated with increased performance in tasks that require focus of attention, ability to accurately discriminate, quick reaction, and retrieval of long-term memory (McCraty et al., 2006). The results of this study lend support to this hypothesis.

Although the data on increased test performance were not as clear in the physiological study sample, there is support for the hypothesis that reducing student stress and text anxiety leads to increased test scores. The fact that there was not a significant increase in overall test scores is not unexpected given that a higher proportion of the control group students in the physiological study sample were from advanced classes (63% compared to 22% in the experimental group) and had much higher tests scores to begin with. However, when baseline test scores were matched on 9th grade CST ELA, there was a marginally significant difference in the test score gains from 9th to 10th grade performance in the experimental group students. Viewed in the light of the overall findings and the smaller number of students in the subgroup (making it harder to achieve statistical significance), we are inclined to regard the marginal result (p = 0.058) for test performance as valid.

Finally, the results from the discriminant function analysis (conducted to investigate the degree to which post-intervention changes on test anxiety and psychophysiological coherence separated the experimental group students from those in the control group), found that before the intervention, both during the resting baseline and the stress preparation periods, the only differentiator of the students in the two groups was student performance on the 9th grade CST ELA test. This result was consistent with the known difference in academic ability between the two schools. However, by the time of post-intervention measurement, test performance was no longer the common differentiator between the two groups of students, but had been replaced by changes in test anxiety and heart rhythm coherence. Not only were these factors effective in discriminating between students in the experimental and control groups during the resting period, but they were an even more powerful discriminator during the stress preparation period—the discriminant function constructed from these factors explained 56% of the variance and achieved a 79% prediction rate in correctly classifying students into their *a priori* appropriate groups.

In short, based on the objective data from electrophysiological measurements, there is a strong pattern of consistent findings from the physiological study showing that the students in the experimental group had learned how to shift into the coherent state and better manage their emotions when preparing for a stressful task or situation, such as taking an important test.

Findings from the Student Drawing Assessment

The drawings from the students in the two schools provided a snapshot of students' experience in the current high-stakes testing era. The drawing assessment was not designed as a clinical psychoanalytical or psychological instrument; rather we looked for patterns and themes that emerged in the drawings as a whole. While coding and analysis of the student drawings is a challenge, due to the emotionally arousing content and the consequent subjectivity of interpretation, that is also the strength of the drawings: they provide a window into the heart and mind of each student as an individual. The student drawings convey what statistics and words alone cannot—the uniquely personal impact of the regime of standardized testing on the individual student.

Two striking findings clearly emerged from a preliminary analysis of the drawings. The first, a surprising finding, is that there were very few depictions of adults in the drawings. Even more startling is the fact that of the few adults shown in the drawings, hardly any are depicted in a positive light. It is likely that this is largely a reflection of the developmental process of adolescent differentiation from adult role models. What is of concern, though, is that the intensely negative experience of test-taking may needlessly exacerbate the marginalization of youth from the adult world.

The second major finding is that the vast majority of the drawings convey intense, mostly negative emotions. As reflected in the quantitative analysis, an overwhelming majority of students (69%) depict negative feelings about the testing experience in their drawings. While there were a few depictions of students with a positive response to tests, these tend to be simpler illustrations, without an elaboration of images, thoughts, or feelings.

The primary negative emotional patterns depicted in these drawings included: feelings of isolation and marginalization; feelings of self-diminishment and insignificance; feelings of fear and of extreme anxiety—specifically test anxiety (students drew themselves biting their nails, sweating profusely, shaking, crying, and pulling their hair out); and feelings of concrete extreme negativity or emptiness—including violent, suicidal, and other destructive or self-destructive gestures and images. The drawings also contain many dark and evocative metaphors—hearts beating out of chests; students trapped as flies in spider webs; students with fire coming out of their mouths and backs, or being engulfed by flames; brains as tornadoes or as gnarled masses of half-done math equations and history dates; students falling or jumping off cliffs; and heads as empty containers, among many others.

Our analysis of the pre–post-intervention changes in the drawings identified patterns of change along three dimensions in the intervention school sub-samples—a movement from a negative to a positive self-concept, a movement from negative to positive feelings and cognitions, and a movement from negative to positive perceptions of self-control and success. In contrast to these clear patterns of change, there were very few instances of positive pre-post changes observed in the student drawings from the control school sub-samples.

The pre-post drawings also suggested that, after participation in the TestEdge program, the intervention school students appear to have developed a greater degree of emotional maturity and increased objectivity both with regard to their feelings and attitudes about testing and also in how they viewed the testing situation. There was also evidence of growth in self-perception and self-awareness, both in moving to more detailed and more human-like depictions of self, and also in moving to depicting a richer, more nuanced expression of different emotions and feelings.

Although these are preliminary findings which await confirmation from the full analysis of the student drawings data currently being conducted, the primary themes are highly consistent with the major findings from the analysis of student perceptions, classroom observations, and physiological measurements. The stark picture that emerges is one of an overwhelming negative reaction that most students have to high-stakes tests and the physiological, mental, and emotional costs directly associated with such high-pressure test-taking.

In these difficult times, this component of the study highlights the significant challenges high-pressure test-taking poses for educators. If such mandated testing is inevitable, it is important for educators to make efforts to mitigate its deleterious effects by providing teachers and students with the tools and techniques to counter the negative, potentially damaging impact of such tests.

Findings from the Qualitative Data

To supplement and provide a broader context for the quantitative data obtained during the study, an observational protocol was utilized to obtain additional measures of the socioemotional climate and interaction patterns both in the classrooms and, more broadly, within the schools. Observations were conducted prior to the implementation of the TestEdge program, during its implementation, and at the end of the study. Teacher and school administrator interviews were also conducted to gather additional data on the teachers' and administrators' experience with the program.

Several distinctive patterns were found during baseline observations in the intervention school at Time 1. First, the mood and degree of positive affect in the classroom environments ranged from extremely negative to extremely positive, along with fluctuating patterns of student responsiveness in the various classrooms. Related to this was an extreme variability in the affective energy expressed by the teachers. There were also high levels of student disengagement across many regular classrooms coupled with an absence of leadership roles among students. In contrast, the control school had much less variability, with more students focused on learning.

In the comparative analysis of the pre–post-intervention observations, the overall preponderance of evidence suggested that more positive changes occurred at the experimental site while more negative changes occurred at the control site. In more specific terms, when compared to their control group classmates, students at the intervention school had reduced levels of fear, frustration, and impulsivity. They also had increased engagement, emotional bonding, humor, persistence, and empathetic listening and understanding. In addition, they appeared to be happier and more peaceful and initiated contact with other students more frequently. There were fewer changes observed in the classrooms at the control school, although there was less teacher rejection or neglect of students and less impulsivity among the students, more student leadership, and increased cooperative group work. Thus, the observational findings regarding the students' interactions and overall affective tone are broadly consistent with the quantitative findings.

Overall, the interviews with the teachers at the primary intervention site, which were conducted after the completion of the study, were positive in regard to the program's content. A number of teachers reported that at least some of the students had become very involved in actively shifting their attitudes. Teachers also reported observing the students exercising more patience and making more of a connection between how they feel and how they perform.

An educational curriculum can only be as effective as its implementation. Thus, the effectiveness of the TestEdge program and the impact of the HeartMath tools could not be adequately tested unless they were effectively implemented. Therefore, one of the objectives of the observational study was to investigate the implementation aspects of the program. The mid-study observations and teacher interviews were both used in the assessment of the implementation process.

The general consensus of the observational team is that in some particulars the TestEdge program was not effectively implemented, indicating that the true potential of the program was not fully realized. The primary reason for this appears to stem from the fact that a number of the teachers resented being told not long before the beginning of school that they had to add an additional program to their teaching load. There is little doubt that this had a negative impact on some of the teachers and affected their motivation and engagement with the program, both in terms of their personal use of the tools and in teaching them to students. The observers reported that some of the teachers were biased against the program before they even knew what it was, some actively resisted, and some were simply resigned and unenthusiastically went through the motions of teaching the tools and concepts. Many felt the program was taking away from their ability to teach their core subject, which they concluded would reflect negatively on them personally if their students did not do well on the mandated tests.

These feelings are understandable and they reflect a sense of the teachers' care about their students. In all fairness, student tests scores do reflect on the teachers and on their school. Fortunately, we were informed that tests scores, as compared to the previous year's class, did increase, although this comparison was not included in the analysis of this controlled study. It was also not entirely fair to the teachers that the study had to be implemented in such a short time frame due to requirements of the grant. These are, however, problems that are easily solvable in the absence of restrictions on funding and time.

Despite this resistance to the program, a number of teachers saw the need for a program like TestEdge and were quite supportive of the process. However, in the final analysis, given that routine practice and sincere engagement are fundamental prerequisites for success of any program such as TestEdge, the lack of an adequate teaching commitment and integration into the classroom in some instances likely negatively impacted the results.

Findings from the Secondary Study

The secondary study was designed to investigate the receptivity, coordination, and administration of the TestEdge program in schools with different cultural, administrative, and situational characteristics. There were a number of important findings from this study. To begin with, there was a universal recognition by teachers that most students are not coming to school emotionally prepared to learn, and that the lack of emotional self-management education for students has resulted in significant blocks to learning. Most teachers involved in the secondary study felt that the time necessary to provide such education would be well worth the investment since it would save considerable time by improving students' ability to learn and perform while diminishing time devoted to managing disruptive behaviors.

Importantly, very few teachers felt that their previous training had given them the skills to deal with their own stress, let alone help their students do so. All of the teachers reported that they personally benefited from the Resilient Educator program and most felt that they would benefit from additional training. The importance of the role of teachers in introducing the intervention cannot be overstated. Teachers who were able to use the tools in their own lives were better able to help students do so. These were also the teachers who reported clear improvements in student behaviors and performance and who felt the program would have long-lasting benefits.

Another conclusion is that teaching two TestEdge lessons per week produced better results than teaching a lesson daily or weekly. This seemed the optimal frequency to allow students time to integrate the lesson material until the techniques become familiar and automatic. Related to the process of internalization, almost all teachers placed a high value on the Freeze-Framer technology, both for themselves and for their students. This was also true at the primary sites where many of the students chose to use the technology on their own, both before and after school.

As with the primary study, the teachers and administrators felt that the implementation process was too abrupt to smoothly integrate the program into the classrooms. This was primarily due to the grant approval being obtained after the school year had already begun. However, even with this limitation, many of the schools were successful in implementing the program and all but one expressed a wish to continue the program. A final conclusion was that the single most important factor in successful implementation was the degree of support for the intervention by the principal and other key school administrators.

Limitations

Study's Implementation

We undertook the TestEdge National Demonstration Study with what we considered a clear vision and high expectations. Based on our extensive research on the effectiveness of HeartMath tools and technology in diverse populations and in a variety of settings in schools and colleges across the nation, we sought to further demonstrate the efficacy of this stress and emotional management technology in a larger-scale study. Our intention was that, with proper implementation, a program designed specifically to counter student stress and test anxiety would result in reduced physiological and emotional manifestations of stress, improved classroom behaviors, enhanced relationships among students and between teachers and students, and, ultimately, improved academic performance.

While the results from this study show that we were partially successful in accomplishing these goals, they also reveal that we were somewhat idealistic in our expectations of schools' and teachers' willingness to embrace innovative solutions for their students. Given the wide publicity regarding test stress in the national media, we anticipated that it would be relatively easy to find schools interested in participating in the study. What we discovered is that so many schools are under so much pressure—particularly to meet state and federal testing requirements—that they could not even take the time to seriously evaluate the opportunity to participate in a U.S. Department of Education-funded project designed to help students relieve test stress and increase test scores. In other instances, schools declined the invitation to participate because they were involved in implementing new curricula, were coming up for review and did not want to compromise their preparations, had a new administrative staff still not familiar enough with the school to feel confident about introducing a new intervention, and so forth.

Originally, we hoped to identify school systems in both Florida and Ohio that would serve as the experimental and control sites. We did considerable research on schools in these states, spoke to educational researchers and administrators to help identify possible sites for the study, and identified matching schools as experimental and control sites. However, time and again, we continually ran up against insurmountable bureaucratic obstacles and/or entangled political systems that ultimately stymied our efforts to recruit suitable sites. For example, within the Pinellas County (FL) school district, after substantial research and, with the help of Congressional staff, identifying schools which not only matched our criteria but who were willing to participate in the study, we ran up against a rigid district policy that forbade any research by non-district personnel. In Ohio, we identified a school district ideal for our research only to discover that they were in the process of restructuring their three high schools into a number of smaller schools, thus making it impossible to meet the population size required for the study. Elsewhere in Ohio, in cities like Cleveland, Toledo, and Akron, we found schools too overwhelmed with immediate problems to even consider entertaining the idea of participating in a major study. We encountered a similar problem with some of the California high schools we contacted.

Even finding secondary study sites proved challenging in many instances. In a high school in Wisconsin, for example, the problem was that the mandated state test was administered in the Fall instead of the Spring, which required us to greatly accelerate our training of observers and teachers. In Pennsylvania, the problem stemmed from the principal deciding to implement the program in a homeroom class in which there was limited time and then using that time for other purposes, which meant that the time for the TestEdge program was significantly diminished.

Unable to find primary sites in Florida and Ohio, we turned to the area proximate to our research center—Northern California. Again, we encountered some of the same problems—schools unwilling or unable (for various reasons) to commit to the study. Although we identified schools that were closely matched in populations and academic performance, we were not successful in recruiting them to participate.

Another problem was relying on the California Department of Education statistics which, while providing data that caused us to identify some schools as well-matched, ultimately produced schools which were quite disparate. While statistics in terms of ethnic make up, state API score ranking, and size made the two schools we identified for the study seem proximate, it turned out that difference in culture, affluence, and other factors, which did not show up in statistical profiles, led to experimental and control sites which were not as similar as we expected they would be. This meant that schools chosen for the study had some, but not all, of the characteristics in common required by our research design. Obviously, this presented some significant unanticipated challenges in gathering, analyzing, and interpreting the data. Although, overall, we believe we have largely met these challenges, they, nevertheless, required flexibility and numerous adjustments and also imposed certain limitations (as discussed below) that we did not originally anticipate.

Perhaps the most sobering revelation stemming from this study is that introducing a highly innovative educational program into a large, entrenched school system and expecting it to take root and be sustained without significant ongoing support in just a few months is unrealistic. Most modern American high schools are large, complex, and often dysfunctional systems. As they have grown in size, increased in ethnic diversity, and been affected by a range of social, economic, and cultural factors, they have become increasingly resistant to change, especially if such change challenges routine patterns of organization and activity, familiar ways and methods of doing things, and, especially, the established authority of those holding positions of power.

Thus, even though our study was well-designed, our intervention strategy well thought-out, and our program innovative; even though we were able to gather significant qualitative and quantitative data; even though we saw significant adoption of HeartMath tools and principles by some teachers and students; ultimately, we were disappointed by the overall resistance, lack of flexibility, and unwillingness to entertain new programs that we encountered in the majority of the schools. This is not say that there were not certain individuals at all levels within the schools who were open, caring, and eager for change, especially in areas that were so desperately needed and acknowledged—there were a number of such individuals, and without their support and facilitation, this study would not have been possible.

We realize that there are many stressors and other factors affecting teachers and students and that it is unrealistic to expect a single, one-time intervention program to affect students' overall academic performance within a period of a few months. It was our original intent to introduce the TestEdge program over a longer period, train all the 10^{th} grade teachers and, in the following year, train the 11^{th} grade teachers so that the practice of the TestEdge tools would be better integrated and sustained. Unfortunately, this was not possible as the grant we received was less than half of the amount requested and the time to complete the project was limited to a single year.

However, based on the findings of this study, we are optimistic that, given enough time and the right conditions, the TestEdge program can be one of the factors that can transform an educational environment, producing improvements not only in behaviors and relationships, but in academic performance as well. Although we hoped to see a more robust increase in test scores in the overall population that participated in the study, we did see a number of subpopulations whose test scores did in fact rise significantly.

In reflecting on the study, even with the challenges we encountered, we learned much about what is required to successfully implement a stress and text anxiety reduction program in a variety of public schools. We also learned a great deal about student attitudes, perceptions, and experiences of stress and test anxiety. We learned more about the sources of students' stress and found that, in general, students want to learn how to self-regulate their emotions to reduce stress. We also found that an innovative program such as TestEdge can be extremely effective in not only reducing students' text anxiety, but also in countering their feelings of anger, loneliness, and disappointment, and in improving their emotional awareness and control and the quality of their social relations. This was an especially encouraging finding in light of the fact that the program was supported by only roughly half of the teachers, there were significant time pressures, and the program was introduced late in the academic year.

Primary Study

For the primary study, the most important limitations are those related to the differences between the intervention and control schools. Although considerable effort was expended to find as closely matched schools as possible, and although the students were closely matched on the majority of factors, including levels of test anxiety, there were two important differences that likely affected our ability to test the hypothesized link between reduced test anxiety and academic performance. The large difference in baseline academic performance and the differences in ethnic composition of the two schools almost certainly worked against the study and likely reduced the effects that would otherwise have occurred in student test performance by reducing stress and anxiety. Although this problem was offset to some degree by statistically controlling for these differences, where possible, it was still an important limitation.

Other important limitations on our ability to fully test the hypotheses were the rushed and therefore less-than-ideal implementation of the TestEdge program and the inability to implement long-term follow-through and measurement. In the short time frame of four months with students' limited exposure to the TestEdge program, it would be unrealistic to expect an overall increase in test scores. The fact that the results showing that the population as whole had reduced test anxiety and exhibited associated improved emotional disposition and interactional effectiveness is quite remarkable. In order to see the full benefit of such a program, students would need more time and facilitation in integrating the concepts and emotional refocusing tools into their daily lives, which would require a longer time period than was possible with this study.

Another limitation that affected our ability to test the study's hypotheses, and therefore the program's true effectiveness, was our inability to measure the primary elements of the hypothesis with adequate specificity and rigor. Postulated in theoretical terms, the hypotheses are stated with the strong assumption that the predicted relationships will hold insofar as "all other things being equal." In actuality, however, the conditions encountered in the study meant that certain key conditions (ethnic composition, socioeconomic and API differences between the two schools) were a lot more unequal than we anticipated because it was not possible to adequately control for these important variables.

Also, the lack of additional information regarding individual differences among students almost certainly limited our ability to fully test the hypothesis and therefore impacted the results. Given the measurement priorities for the *Student Opinion Survey* questionnaire, we did not have the space to determine which students were motivated to do well on the tests, their levels of preparation, their facility with English, and so forth. It is likely that the higher percentage of students at the experimental school for whom English was not their primary language impacted the English test scores, and likely impacted the Math test scores as well. We had inadequate data for clearly identifying students for whom English is a second language. Even though it was readily apparent from the ethnic breakdown of student enrollment in the two schools that the intervention school had a notably higher proportion of students for whom English is not their primary language (surprisingly, in comparison to the control school, a disproportionate percentage of students at the intervention school were enrolled in ESL classes), we were unable to obtain specific information at the level of the individual student to construct a standardized measure for all students in the two schools. Thus while it is highly likely that this difference in English language proficiency had a significant impact on the CST English-Language-Arts test scores, without a standardized measure across the two schools, we were unable to investigate the nature and consequences of this difference.

A final limitation was the lack of uniformity in the standardized tests at the two schools. This severely limited our ability to adequately compare the groups in term of overall test performance, although this was overcome to some degree by using statistically-matched subgroup comparisons.

There were some limitations in the physiological study. First we were not able to measure the physiological parameters while the students were actually taking the high-stakes tests; rather, we had to simulate the testing environment using the computerized Stroop Color-Word Conflict Test. It is possible that had we been able to measure students' physiological processes during an actual high-stakes test, we may have found different results. It is more likely, however, given the stress and anxiety that most students report experiencing when taking an important test, that a stronger relationship between the physiological parameters and test performance would have been found.

While we hope to address these limitations in future studies, this study has provided new and important findings regarding the interactions between physiology, emotions, learning, and performance. In our analysis of the relationships between test anxiety, coherence, and test performance, there appear to be associations between increased psychophysiological stability and academic performance. The data suggest that when students self-manage their stress using coherence-building tools, it enables them to achieve both a significant reduction in testing-related anxiety and a corresponding improvement in standardized test scores.

The Question of Causal Inference

To determine whether or not the TestEdge intervention was successful involves evaluating the study's results in relation to the question of causal inference: assessing the degree to which the intervention *actually caused* changes in the internal processing and behavior of students in the experimental school that resulted in observed outcomes consistent with the study's theoretical expectations. As will become clear in the next few paragraphs, this is a complex issue that is not easily resolvable, even under the more ideal circumstances of a controlled laboratory experiment—let alone when dealing with an open field study like this one.

To adequately address the measurement requirements for causal inference, the study utilized a quasi-experimental design in which one school was treated as the treatment group—into which the TestEdge program was introduced—and the second school was treated as a control. The operational requirements for causal inference entailed pre- and post-measurements on each side of the TestEdge intervention in the experimental group: one measurement at baseline before the TestEdge intervention was implemented, and a second measurement, using the same instrumentation, at the end of the treatment period, just before the students took the CST.[34] While this research design meets Campbell and Stanley's (1963) criteria for causal inference in a quasi-experimental design, there are some additional considerations that require the exercise of caution when weighing evidence of causality in studies that utilize a two-panel repeated measures design, such as our study.

According to Stinchcombe (1968), the empirical requirements for establishing an adequate basis for causal inference demands, at minimum, pre- and post-measures and showing that the overtime variance of the dependent variable is directly associated with the overtime change (in the expected direction) of the independent variable. Our study design meets this basic requirement.

However, in this research design a key condition for establishing causality is missing: namely, the *processes by which* the causal effect from the independent variable is actually transmitted to and produces a change in the behavior of the dependent variable (Bradley, 1987). To establish the empirical requirements for inference of causal process requires continuous measurement of both the dependent and independent variables, and, insofar as the effect of the independent variable is transmitted by some other intermediary or intervening factor/s, continuous measurement of those factors

[34]It will be recalled that the test anxiety instrument was administered in a four panel format: in addition to pre- and post-measurement, the test anxiety instrument was administered two additional times—one week before students took the CAHSEE in March and one week prior to the CST in April.

as well. Such requirements usually can only be met in the controlled conditions of a laboratory experiment, which is what we endeavored to achieve in the design of our physiology study. In terms of the experimental protocol for the physiology study, we had continuous measurement of each student's psychophysiological state, via electrophysiological instrumentation, and continuous measurement of each student's test performance, via his/her performance of the set of real-time tasks in the Stroop Test.

These considerations lead to the following conclusions. First, even assuming optimal conditions for our primary study, the most that can be inferred in terms of the causal effects of the TestEdge intervention is that there is evidence of covariance in test anxiety and test performance consistent with the expected effects of the intervention. This means that even though we have compelling evidence of such covariance in our matched-groups analysis sub-samples, we do not have the requisite empirical basis for causal inference in the *strong* sense of actually having empirically established the causal processes by which the intervention brought about changes in test anxiety and test performance.

Second, although we are closer to achieving measurement of causal process in the physiology study, two caveats must be borne in mind when assessing the study's results. One is that we did not have comprehensive measurement of *all* of the bodily processes involved in stress response and task performance. While the HRV measures we used likely captured important bodily processes—namely, psychophysiological processing of emotional arousal—due to practical considerations, we were unable to measure other basic elements, such as cognitive processing and neurological activity. Also, while we used the Stroop Test to simulate the stressful conditions of taking a standardized academic test, the Stroop Test is not an achievement test and, therefore, consideration must be given to how closely it approximates a student's actual experience of taking the CAHSEE or the CST.

The Question of Generalizability

To this point, we have discussed our major findings and reviewed the study's limitations. And while we believe, in evaluating the evidence and weighing the impact of the limitations, that our primary findings are valid, a major question regarding the study's broader significance concerns the question of its generalizability—the degree to which the study's primary findings are valid and applicable to schools in America in general.

This issue can be addressed in two ways. The first is to perform certain kinds of statistical analyses that enable an assessment of this question on statistical grounds. The approach that we will take is to perform a series of split-half reliability analyses, where case counts permit, to evaluate the statistical generalizability of the study's major findings. In a typical application, the generalizability of the results from a given statistical model are subjected to a further verification by repeating exactly the same analysis procedures on a fifty-percent random sample of the population upon which the original analysis was conducted. The results from the split-half sample are then compared with the original results for evidence that they are statistically comparable. Because this randomization procedure is statistically the equivalent to replication of the study on an independent random sample, this provides the investigator with a means of estimating the likely generalizability of a given result. As we are still in the process of completing the analysis of all of the data in the study, we have not yet had the opportunity to conduct these generalizability analyses. These will be conducted once the analysis is complete and before the results of the study are submitted for journal publication.

A second approach to the question of generalizability is an evaluation of the degree to which the study population—the samples of students from the two high schools—are representative of high school populations in California and, more broadly, of high schools across the nation. While this question is ultimately resolved by additional studies in schools in California and elsewhere, which would aim to test the generalizability of our findings by replicating the design and implementation of our research, some initial evaluation is possible with the data at hand.

To the degree that schools and California and elsewhere are comparable to the experimental school in our study in terms of context and school size and type, and also in socio-ethno-economic composition of the student population, there is a high likelihood that the effects of the TestEdge program observed in our study will also hold for these other schools. Moreover, insofar as there are improvements in school and teacher buy-in to the program, the effects of the intervention are likely to be enhanced and the expected outcomes even stronger.

The question of generalizability becomes more difficult in schools and student populations where differences in context, size and type of school and student composition may complicate the likely effects of the program and the expected consequent outcomes. Even so, while it is clear from our results that the incidence of test anxiety varies by gender and ethnicity and, to a lesser degree, by family composition and class academic level of the student, it is also clear that students across *all* of these characteristics suffer from test anxiety and are handicapped by its negative effects on their test performance. Moreover, it is also clear that the overwhelming majority of students across most of these characteristics[35] gained a clear benefit of the TestEdge program in their reduced level of test anxiety and, in sizeable numbers in certain sub-samples, also benefited from an associated improvement in their test scores. Thus, insofar as test anxiety is a ubiquitous phenomenon that can deleteriously affect virtually any student in any school, and insofar as the TestEdge tools appear able to be learned and used by virtually any student at any level of academic ability, then it would seem that there is no good reason why the findings documenting the benefits of the TestEdge program in this study should not also be expected for most students in other school settings. One important caveat to this conclusion is that the program be implemented in a bottom-up approach with enthusiastic support by the school administration and teachers.

[35] The notable exception is male students who were less likely to evidence a significant reduction in test anxiety and also less likely to achieve an associated improvement in test scores.

Chapter XVI
Conclusions and Implications

The more educators come to understand the nature of the relationship between emotion and cognition, the better they may be able to leverage this relationship in the design of learning environments.

—Mary Helen Immordino-Yang & Antonio Damasio (2007:9)

In addition to presenting the primary conclusions from this research, we also want to highlight some important implications of our major findings for the scientific understanding of student test anxiety, learning, and academic performance.

Recapitulation

The TestEdge program was developed by the Institute of HeartMath as a supplementary program to help students reduce stress and test anxiety, improve test and academic performance, and improve the emotional and relational dynamics of the classroom and school. Based on 15 years of scientific research on the physiology of emotions and heart–brain communication, this program employs a set of easy-to-use, positive emotion-refocusing and restructuring techniques that enable students to better recognize and self-regulate stress, test anxiety and other emotion-based blocks to learning and test performance. A key element of the effectiveness of the TestEdge techniques is the self activation of a specific, scientifically measurable physiological state associated with increased synchronization in nervous system activity, increased emotional stability, and improved cognitive performance. Although the TestEdge program has been successfully implemented in schools throughout the U.S. and in some foreign countries, and although pilot studies have shown improvements in standardized test scores and passing rates, until now it had not been tested in a large-scale implementation.

This Department of Education-funded study was therefore initiated to determine the impact of this learning program on larger student populations and in diverse school settings. The study's primary objectives were: 1) to determine the nature, magnitude, correlates, and consequences of stress and test anxiety among students; and 2) to investigate the efficacy of the TestEdge program in improving student test anxiety, test performance, emotional well-being, quality of relationships, and classroom dynamics.

By analyzing a combination of physiological, psychological, emotional, and sociological data collected within a multi-methods framework involving quantitative and qualitative measurement and methods of analysis, this comprehensive investigation generated a number of important findings.

Our study found that 61% of our sample of tenth-grade students was affected by test anxiety. Furthermore, 26% of the sample suffered from high levels of test anxiety. We also found that twice as many girls as boys are handicapped by a high level of test anxiety. The study confirmed previous findings of a link between high test anxiety and reduced academic performance. On average, students with high levels of anxiety scored 15 points lower on both Math and English-Language Arts standardized tests than those with low test anxiety.

We also found that student test scores were associated with their feelings of connection to one another, their sense of safety and happiness at school, their perceptions of how much school would help them in getting a job or going to college, and their feelings of having a purpose in life. Within this context, the challenge for educators is how best to prepare students to deal with the increased stress they face so that their test performance is reflective of their true ability.

A major finding from this study is that the majority of students are affected by test anxiety. Moreover, without the use of formal instruments to gauge this anxiety or effective interventions to counter it, it is unlikely that educators have an accurate picture of a student's true academic potential. To provide educators with such information, it is necessary to deepen our understanding of student test anxiety and to implement effective tools to help students manage their emotional stress when taking important tests.

Relative to this issue, an important finding in this study was that the TestEdge program was highly successful in producing significantly lower mean levels of test anxiety in the general population of students who participated in the program. This reduction in test anxiety occurred right across the spectrum of academic ability. In addition, students had significant reductions in feelings of loneliness, sadness, anger, depression, and disappointment, and significant improvements in their ability to be aware of and control their emotions. They also had significant improvements in positive emotions and feelings, in their awareness of others' feelings (empathy), in their ability to share their feelings with others, in their sense of being understood, and in their ability to avoid conflicts.

Despite sampling differences in API, ethnic composition, and socio-economic status favoring the control school over the intervention school, which likely masked the effects of the TestEdge intervention, we still found compelling evidence of a reduction in test anxiety in association with an improvement in test performance in the intervention school. Using a matched-groups approach, where we could better control for differences between the groups, significant increases in test scores were consistently found in subgroups of students in the experimental group, indicating that the TestEdge program is effective in helping students manage test anxiety and increase test performance. Here the results showed that female students appeared to benefit the most.

The physiological study provided support for a causal link between increased psychophysiological coherence and cognitive functions critical in learning and test-taking. It also provided an objective measure which convincingly demonstrated that the students had acquired the ability to shift into the coherent state prior to taking an important test. Beyond this, the physiological data indicated that through their consistent use of the HeartMath tools over the study period, the students in the TestEdge program had instantiated a healthier, more harmonious, and more adaptive baseline pattern of psychophysiological function: they exhibited increased HRV and heart rhythm coherence even without conscious use of the tools. This is a particularly significant finding as it provides evidence of the occurrence of a repatterning process at a fundamental level, whereby coherence is progressively established in the neural architecture as a new, stable psychophysiological baseline or set point. When neurologically instantiated, the result of this systemic repatterning is a sustained improvement in the patterns of psychophysiological activity underlying emotional experience, emotion regulation, stress resiliency, and cognitive function.

In short, we can conclude that the consistency across the various quantitative, physiological, and observational findings demonstrates that the TestEdge program led to a number of important successes. There is good reason to believe that the program generated substantial benefit in the physiological, psychological, academic, and social dimensions of students' lives.

Implications

There are two kinds of implications this study has for research on test anxiety and test performance—methodological and theoretical. Our use of a diverse combination of measurement modes and methods stands in contrast with the traditional approach used in most education studies where investigators typically choose either a quantitative or qualitative approach to their research. In addition to the use of standardized tests, questionnaires, interviews, and school and classroom observations, we also added community profiles, a student drawings assessment, and an electrophysiological sub-study. Going well beyond the narrower approach normally employed, we believe this kind of rich combination of methods and procedures not only provides a more robust foundation of cross-corroborating evidence upon which to base our major findings, but this approach also leads to a number of new and substantively informative findings, such as those from the student drawings assessment and the physiological study. In short, we believe education research has much to gain by complementing standard methodologies with the more novel methods and procedures utilized in this study.

Despite the enormous time, effort, and resources involved, the electrophysiological recordings added an entirely new window on student emotions and test anxiety. Not only is this a window to a new level of analysis—namely, that of the psychophysiological level—but the data collected are objective and provide a description of the physiological substratum of stress and test anxiety that is not filtered or distorted by the subjective reality of a student's cognitive constructions or psychological defenses. On the basis of the bounty of findings and potential new understandings offered by the physiological data, we believe the inclusion of physiological measurements holds great promise for deepening the understanding of the relationship of between student emotional states and learning and academic performance.

Moving to the theoretical implications of the study, we believe that the approach and findings from this research point to the value of a significantly broadened perspective on student development, learning, and academic performance and are also in line with the emergence of a new understanding. For some time now, the primary emphasis of education research has been on cognition—memory, intellectual ability, learning skills, comprehension, reasoning, and so forth. However, the major findings from this study add to the growing body of evidence questioning the efficacy of approaches to education that place primary emphasis on cognition.

In more specific terms, the results from this study show that about twenty-five percent of the variance in student test performance is explained by measures of students' emotional states. Notably, the results show that negative emotions have a negative impact on test performance, whereas positive emotions have a positive impact on test performance. In addition, the measures of interactions and relations among students were also related to test performance. The theoretical import of these findings can be seen when it is recalled that the cognition-based measures of test anxiety (Spielberger's Test Anxiety Inventory) only accounted for about twelve percent of the variance in test performance.

Overall, the main thrust of the findings suggests a theoretical perspective in which emotions and relations are key factors in facilitating or inhibiting student learning and performance. This is not to say that cognitive factors are not important. Rather, the point here is to rebalance the long-standing priority given to the cognitive perspective on education. We believe the findings of this study point to the urgent need to develop a new science of education in which developing an understanding of the influence of affective and relational factors on cognition and behavior is the key focus. An important element of such an effort is the need to develop a greater understanding of how the socioemotional dynamics of the classroom and the school affect student learning and academic performance.

A final implication of note concerns the primary finding of the study: namely, that the TestEdge intervention was successful in reducing mean test anxiety overall and also, for certain matched groupings of students, was associated with an improvement in test performance. Given that the TestEdge program's core intervention utilizes a set of positive emotion-based techniques that engage the *whole* psychophysiological system, this finding, too, challenges many of the assumptions underlying cognition-based interventions.

In short, there is much in the findings from this study that not only questions the primacy given to the cognitive perspective on education, but also points to the strategic import of a perspective in which emotions and relations are central. This echoes the critical point made by Immordino-Yang and Damasio (2007: 3) in the epigram at the outset of this report, which is worth repeating here:

> *Recent advances in neuroscience are highlighting connections between emotion, social functioning, and decision making that have the potential to revolutionize our understanding of the role of affect in education. In particular, the neurobiological evidence suggests that the aspects of cognition that we recruit most heavily in schools, namely learning, attention, memory, decision making, and social functioning, are both profoundly affected by and subsumed within the processes of emotion... .*

This new science has enormous practical implications for how we go about designing the social and emotional environments of our schools to facilitate learning and psychosocial growth. As Immordino-Yang and Damasio (2007: 9) go on to say, as we continue to deepen our understanding of this fundamental relationship between emotion and cognition, we can more effectively "leverage this relationship" to create optimal educational environments in which students will flourish.

Conclusion

It is our hope that the results of this research will not only make an important contribution to the scientific understanding of the emotional and psychophysiological factors affecting test anxiety and test performance, but that they will also have an important impact on educational policies in schools. It is clear that we cannot continue to ignore students' experience of emotional stress and its impact on their academic and psychosocial development. By introducing and sustaining appropriate stress and emotional self-management programs for students of all ages and their teachers, it should be possible to significantly reduce the emotional discord that impedes learning and performance. The integration of such programs into educational curricula on a large scale will increase the effectiveness of our educational system and, in the long-term, help to boost the academic standing of the United States in the international community.

References

Arguelles, L., McCraty, R., & Rees, R. A. (2003). The heart in holistic education. *Encounter: Education for Meaning and Social Justice, 16*(3), 13-21.

Arguelles, L. & Schroeder, L. (2004). *Classroom Observation Protocol*. Unpublished document, TestEdge National Demonstration Study Project, Claremont Graduate University, Claremont, CA.

Arguelles, L., Schroeder, L., Daugherty, A., & Hollingsworth, T. (2006). *Executive summary report: Inquiry work conducted by the CGU Research Team for the TestEdge National Demonstration Study*. Unpublished document, Claremont Graduate University, Claremont, CA.

Armour, J. A., & Kember, G. C. (2004). Cardiac sensory neurons. In J. A. Armour & J. L. Ardell (Eds.), *Basic and Clinical Neurocardiology* (pp. 79-117). New York: Oxford University Press.

Aspinwall, L. G. (1998). Rethinking the role of positive affect in self-regulation. *Motivation and Emotion, 22*(1), 1-32.

Barlow, D. H. (1988). *Anxiety and Its Disorders*. New York: Guilford Press.

Barrios-Choplin, B., McCraty, R., & Cryer, B. (1997). An inner quality approach to reducing stress and improving physical and emotional wellbeing at work. *Stress Medicine, 13*(3), 193-201.

Beck, A. T. (1976). *Cognitive Therapy and the Emotional Disorders*. New York: International Universities Press.

Beidel, D. C. (1991). Social phobia and overanxious disorder in school-age children. *Journal of the American Academy of Child and Adolescent Psychiatry, 30*, 545-552.

Beidel, D. C., Turner, S. M., & Taylor-Ferreira, J. C. (1999). Teaching study skills and test-taking strategies to elementary school students. The Testbusters Program. *Behavior Modification, 23*(4), 630-646.

Benson, H. (1975). *The Relaxation Response*. New York: Morrow Press.

Bouman, S. (2006). Unpublished data, Claremont Graduate University, Claremont, CA.

Bradley, R. T. (1987). *Charisma and Social Structure: A Study of Love and Power, Wholeness and Transformation*. New York: Paragon House.

Bradley, R. T. (2001). The origin and organization of socioaffective dialogue. International Journal of Public Administration, 24(7-8), 799-842.

Bradley, R. T. (2004). Love, power, brain, mind, and agency. In D. Loye (Ed.), The Great Adventure: Toward a Fully Human Theory of Evolution (pp. 99-150). Albany: State University of New York Press.

Bradley, R. T., & Atkinson, M. (2004a). Classroom Relations Survey. Boulder Creek, CA: Institute of HeartMath.

Bradley, R. T., & Atkinson, M. (2004b). Student Opinion Survey. Boulder Creek, CA: Institute of HeartMath.

Bradley, R. T., & Atkinson, M. (2004c). Teacher Survey. Boulder Creek, CA: Institute of HeartMath.

Bradley, R. T., Young, W. Y., Ebbs, P., & Martin, J. (1993). Characteristics of advanced-level dietetics practice: A model and empirical results. Journal of The American Dietetic Association, 93(2): 196-202.

Bradley, R. T., Goelitiz, J., Atkinson, M., & Tomasino, D. (2004). TestEdge Assessment. Boulder Creek, CA: Institute of HeartMath.

Bradley, R. T., McCraty, R., & Rees, R. (2004). Proposal concept for a study of highly effective and transformational teaching. Unpublished manuscript, Institute of HeartMath, Boulder Creek, CA.

Bradley, R. T., & Pribram, K. H. (1998). Communication and stability in social collectives. Journal of Social and Evolutionary Systems, 21(1), 29-80.

Bush, K. J. (1989). Test Anxiety in Second, Third, and Fourth-Grade Students: An Investigation of Group Differences, Teacher Anxiety, Achievement, and Classroom Attitudes. Doctoral dissertation, Georgia State University, Publication No. AAT 8922902.

Cameron, O. G. (2002). Visceral Sensory Neuroscience: Interoception. New York: Oxford University Press.

Campbell, D. T., & Stanley, J. C. (1963). Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally College Publishing Company.

Cannon, W. B. (1929). Bodily Changes in Pain, Hunger, Fear and Rage: An Account of Recent Researches into the Function of Emotional Excitement, 2nd edition. New York: D. Appleton & Company.

Cantin, M., & Genest, J. (1986). The heart as an endocrine gland. Scientific American, 254(2), 76-81.

Childre, D., & Martin, H. (1999). The HeartMath Solution. San Francisco: Harper San Francisco.

Childre, D., & Rozman, D. (2005). Transforming Stress. Oakland, CA: New Harbinger.

Childre, D., & Rozman, D. (2006). Transforming Anxiety. Oakland, CA: New Harbinger.

Cizek, G. J., & Burg, S. S. (2006). *Addressing Test Anxiety in a High-Stakes Environment: Strategies for Classrooms and Schools*. Thousand Oaks, CA: Corwin Press.

Clark, K. B., Naritoku, D. K., Smith, D. C., Browning, R. A., & Jensen, R. A. (1999). Enhanced recognition memory following vagus nerve stimulation in human subjects. *Nature Neuroscience, 2*(1), 94-98.

Cohen, J. (2006). Social, emotional, ethical and academic education: Creating a climate for learning, participation in democracy, and well-being. *Harvard Educational Review, 76*(2), 201-237.

Costa, A. L., & Kallick, B. (2000). *Habits of Mind: A Developmental Series*. Alexandria, VA: Association for Supervision and Curriculum Development.

Damasio, A. (2003). *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain*. Orlando: Harcourt.

Daugherty, A. K. (2006). *Physiological, Cognitive, and Psychosocial Effects of Emotional Refocusing: A Summative and Formative Analysis*. Doctoral dissertation, Faculty of Education, Claremont Graduate University, Claremont, CA.

Drew, D. E. (1996). *Aptitude Revisited: Rethinking Math and Science Education for America's Next Century*. Baltimore: Johns Hopkins University Press.

Dworkin, J. B., Larson, R., & Hansen, D. (2003). Adolescents' accounts of growth experiences in youth activities. *Journal of Youth and Adolescence, 32*(1), 17-26.

Edelstein, D. (2000). Tests + Stress = Problems for Students. *Brain Connection:* www.brainconnection.com/topics/?main=fa/test-stress. Date Accessed: 10/17/2006.

Elkind, D. (2001). *The Hurried Child*. Cambridge: Perseus Publishing.

Erford, B. T., & Moore-Thomas, C. (2004). Testing FAQ: How to answer questions parents frequently ask about testing. In J. E. Wall & G. R. Walz (Eds.), *Measuring up: Assessment Issues for Teachers, Counselors, and Administrators* (pp. 535-555). Greensboro, NC: ERIC Clearinghouse on Counseling and Student Services.

Evers-Lush, M. J. (1991). *Factors, Including Curriculum, That Relate to Test Anxiety Experienced by Secondary Students: A Study Based on High School and Beyond.*

Doctoral dissertation, Virginia Polytechnic Institute and State University, Publication No. AAT 9218506.

Fanelli, R. M. (1991). *Cognitive Assessment of High School Students' Test Anxiety Regarding SAT Performance (High School Students)*: Doctoral dissertation, New York University, Publication No. AAT 9124568.

Fredrickson, B. L. (2002). Positive emotions. In C. R. Snyder & S. J. Lopez (Eds.), *Handbook of Positive Psychology* (pp. 120-134). New York: Oxford University Press.

Friedman, B. H. (in preparation). An autonomic flexibility-neurovisceral integration model of anxiety and cardiac vagal tone.

Friedman, B. H., & Thayer, J. F. (1998a). Anxiety and autonomic flexibility: A cardiovascular approach. *Biological Psychology, 49*(3), 303-323.

Friedman, B. H., & Thayer, J. F. (1998b). Autonomic balance revisited: Panic anxiety and heart rate variability. *Journal of Psychosomatic Research, 44*(1), 133-151.

Goleman, D. (1995). *Emotional Intelligence*. New York: Bantam Books.

Greenberg, M. T., Weissberg, R. P., O'Brien, M. U., Zins, J. E., Fredericks, L., Resnik, H., & Elias, M. J. (2003). Enhancing school-based prevention and youth development through coordinated social, emotional, and academic learning. *American Psychologist, 58*(6-7), 466-474.

Greenberger, D., & Padesky, C. (1995). *Mind Over Mood: Change How You Feel by Changing the Way You Think*. New York: The Guilford Press.

Hargreaves, A. (2004). Inclusive and exclusive educational change: Emotional responses from teachers and implications for leadership. *School Leadership and Management, 24*(2), 287-309.

Hartnett-Edwards, K. (2006). *The Social Psychology and Physiology of Reading/Language Arts*. Doctoral dissertation, Faculty of Education, Claremont Graduate University, Claremont, CA.

Hasebe-Ludt, E. (1999). The whole world in one classroom: Teaching across traditions in a cosmopolitan environment. *Interchange, 30*(1), 35-39.

Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. *Review of Educational Research, 58*(1), 47-77.

Hill, K. T. (1984). Debilitating motivation and testing: A major educational problem, possible solutions, and policy applications. In R. Ames & C. Ames (Eds.), *Research on Motivation in Education: Student Motivation* (pp. 245-274). Orlando, FL: Academic.

Hinde, R. A. (1992). Developmental psychology in the context of other behavioral sciences. *Developmental Psychology, 28*(6), 1018-1029.

Hollingsworth, T. (2006). Observational changes in tenth grade classroom environments: Qualitative measures of *TestEdge* intervention on test anxiety. Unpublished paper, Department of Educational Studies, Claremont Graduate University, Claremont, CA.

Hollingsworth, T. (2007). *Four Years into No Child Left Behind: A Profile of California Tenth Grade Students' Well-Being and Mixed Methods Analysis of Its Associations with Academic Success*. Doctoral dissertation, Faculty of Education, Claremont Graduate University, Claremont, CA.

James, W. (1884). What is an emotion? *Mind, 9*(34), 188-205.

Lagozzino, J. S. (2006). Background information and observations on the implementation of the TestEdge curriculum in third grade at [Southern California] Elementary. Unpublished paper, Claremont Graduate University, Claremont, CA.

Immordino-Yang, M. H., & Damasio, A. (2007). We feel, therefore we learn: The relevance of affective and social neuroscience to education. *Mind, Brain, and Education, 1*(1), 3-10.

Institute of HeartMath (2004). *TestEdge: Getting in Sync for Test Success*. Boulder Creek, CA: HeartMath LLC.

Isen, A. M. (1987). Positive affect, cognitive processes, and social behavior. In L. Berkovitz (Ed.), *Advances in Experimental Social Psychology,* Vol. 20 (pp. 203-253).

Isen, A. M. (1999). Positive affect. In T. Dalgleish & M. Power (Eds.), *Handbook of Cognition and Emotion* (pp. 522-539). New York: John Wiley & Sons.

Isen, A. M., Daubman, K. A., & Nowicki, G. P. (1987). Positive affect facilitates creative problem solving. *Journal of Personality and Social Psychology, 52*(6), 1122-1131.

Lane, R. D., Reiman, E. M., Ahern, G. L., & Thayer, J. F. (2001). Activity in medial prefrontal cortex correlates with vagal component of heart rate variability during emotion. *Brain and Cognition, 47*, 97-100.

Lazarus, R. S. (1966). *Psychological Stress and the Coping Process*. New York: McGraw Hill.

LeDoux, J. (1996). *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. New York: Simon and Schuster.

LeDoux, J. E. (1994). Cognitive-emotional interactions in the brain. In P. Ekman & R. J. Davidson (Eds.), *The Nature of Emotion: Fundamental Questions* (pp. 216-223). New York: Oxford University Press.

Lifford, J., Byron, B., Eckblad, J., & Ziemian, C. (2000). Reading, responding, reflection. *English Journal, 89*(4), 46-57.

Luskin, F., Reitz, M., Newell, K., Quinn, T. G., & Haskell, W. (2002). A controlled pilot study of stress management training of elderly patients with congestive heart failure. *Preventive Cardiology, 5*(4), 168-172, 176.

MacLeod, A. K., & Moore, R. (2000). Positive thinking revisited: Positive cognitions, well-being and mental health. *Clinical Psychology and Psychotherapy, 7*(1), 1-10.

Malchiodi, C. A. (1998). *Understanding Children's Drawings*. New York: The Guilford Press.

McCraty, R. (2003). *Heart-Brain Neurodynamics: The Making of Emotions*. Boulder Creek, CA: HeartMath Research Center, Institute of HeartMath, Publication No. 03-015. Available as an electronic monograph at: www.heartmath.org/research/e-books.

McCraty, R. (2005). Enhancing emotional, social, and academic learning with heart rhythm coherence feedback. *Biofeedback, 33*(4), 130-134.

McCraty, R., Atkinson, M., Tiller, W. A., Rein, G., & Watkins, A. D. (1995). The effects of emotions on short-term power spectrum analysis of heart rate variability. *American Journal of Cardiology, 76*(14), 1089-1093.

McCraty, R., Atkinson, M., & Tomasino, D. (2001). *Science of the Heart: Exploring the Role of the Heart in Human Performance*. Boulder Creek, CA: HeartMath Research Center, Institute of HeartMath, Publication No. 01-001. Available as an electronic monograph at: www.heartmath.org/research/e-books.

McCraty, R., Atkinson, M., & Tomasino, D. (2003). Impact of a workplace stress reduction program on blood pressure and emotional health in hypertensive employees. *Journal of Alternative and Complementary Medicine, 9*(3), 355-369.

McCraty, R., Atkinson, M., Tomasino, D., & Bradley, R. T. (2006). *The Coherent Heart: Heart-Brain Interactions, Psychophysiological Coherence, and the Emergence of System-Wide Order*. Boulder Creek, CA: HeartMath Research Center, Institute of HeartMath, Publication No. 06-022. Available as an electronic monograph at: www.heartmath.org/research/e-books.

McCraty, R., Atkinson, M., Tomasino, D., Goelitz, J., & Mayrovitz, H. N. (1999). The impact of an emotional self-management skills course on psychosocial functioning and autonomic recovery to stress in middle school children. *Integrative Physiological and Behavioral Science, 34*(4), 246-268.

McCraty, R., Barrios-Choplin, B., Rozman, D., Atkinson, M., & Watkins, A. D. (1998). The impact of a new emotional self-management program on stress, emotions, heart rate variability, DHEA and cortisol. *Integrative Physiological and Behavioral Science, 33*(2), 151-170.

McCraty, R., Bradley, R. T., & Tomasino, D. (2004–2005). The resonant heart. *Shift: At the Frontiers of Consciousness*, Special Issue—"The Science of Fields," Dec. 2004–Feb. 2005 (No. 5), 15-19.

McCraty, R., & Childre, D. (2004). The grateful heart: The psychophysiology of appreciation. In R. A. Emmons & M. E. McCullough (Eds.), *The Psychology of Gratitude* (pp. 230-255). New York: Oxford University Press.

McCraty, R., & Tomasino, D. (2006). Emotional stress, positive emotions, and psychophysiological coherence. In B. B. Arnetz & R. Ekman (Eds.), *Stress in Health and Disease* (pp. 360-383). Weinheim, Germany: Wiley-VCH.

McCraty, R., Tomasino, D., Atkinson, M., Aasen, P., & Thurik, S. J. (2000). *Improving Test-Taking Skills and Academic Performance in High School Students Using HeartMath Learning Enhancement Tools*. Boulder Creek, CA: HeartMath Research Center, Institute of HeartMath, Publication No. 00-010. Available at: www.heartmath.org/research/research-publications.html

Mezzacappa, E., Tremblay, R. E., Kindlon, D., Saul, J. P., Arseneault, L., Seguin, J., Pihl, R. O., & Earls, F. (1997). Anxiety, antisocial behavior, and heart rate regulation in adolescent males. *Journal of Child Psychology and Psychiatry, 38*(4), 457-469.

Monk, C., Kovelenko, P., Ellman, L. M., Sloan, R. P., Bagiella, E., Gorman, J. M., & Pine, D. S. (2001). Enhanced stress reactivity in paediatric anxiety disorders: implications for future cardiovascular health. *International Journal of Neuropsychopharmacology, 4*(2), 199-206.

Most, S. B., Scholl, B. J., Clifford, E. R., & Simons, D. J. (2005). What you see is what you set: Sustained inattentional blindness and the capture of awareness. *Psychological Review, 112*(1), 217-242.

Mullen, C. (2002). Building an arts-based curriculum: A lesson from the bower bird. *Teacher Education Quarterly, Fall*, 3-15.

Niedenthal, P., & Kitayama, S. (Eds.). (1994). *The Heart's Eye: Emotional Influences in Perception and Attention*. San Diego: Academic Press.

National Institute of Mental Health. (2000). *Cognitive Research at the National Institute of Mental Health*. Bethesda, MD: National Institute of Mental Health.

Piaget, J. (1965/1995). Logical operations and social life. In L. Smith (Ed.), *Sociological Studies* (pp. 134-157). London: Routledge.

Pine, D. S., Wasserman, G. A., Miller, L., Coplan, J. D., Bagiella, E., Kovelenku, P., Myers, M. M., & Sloan, R. P. (1998). Heart period variability and psychopathology in urban boys at risk for delinquency. *Psychophysiology, 35*(5), 521-529.

Porges, S. W. (1984). Physiologic correlates of attention: A core process underlying learning disorders. *Pediatric Clinics of North America, 31*(2), 371-385.

Porges, S. W. (1992a). Autonomic regulation and attention. In H. Hayne & R. Richardson (Eds.), *Attention and Information Processing in Infants and Adults* (pp. 201-223). Hillsdale, NJ: Erlbaum.

Porges, S. W. (1992b). Vagal tone: A physiologic marker of stress vulnerability. *Pediatrics, 90*(3 Pt 2), 498-504.

Porges, S. W., Doussard-Roosevelt, J. A., & Maiti, A. K. (1994). Vagal tone and the physiological regulation of emotion. In N. A. Fox (Ed.), *Emotion Regulation: Behavioral and Biological Considerations*. *Monographs of the Society for Research in Child Development*, Vol. 59(2-3), pp. 167-186, 250-283.

Pribram, K. H. (1967). The new neurology and the biology of emotion: A structural approach. *American Psychologist, 22*(10), 830-838.

Pribram, K. H. (1984). Emotions: A neurobehavioral analysis. In K. R. Scherer & P. Ekman (Eds.), *Approaches to Emotion*. Hillsdale, NJ: Erlbaum.

Pribram, K. H. (1991). *Brain and Perception: Holonomy and Structure in Figural Processing*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Pribram, K. H., & Melges, F. T. (1969). Psychophysiological basis of emotion. In P. J. Vinken & G. W. Bruyn (Eds.), *Handbook of Clinical Neurology* (Vol. 3, pp. 316-341). Amsterdam: North-Holland Publishing Company.

Ratey, J. J. (2001). *A User's Guide to the Brain: Perception, Attention, and the Four Theaters of the Brain*. New York: Pantheon Books.

Richards, J. E. (1987). Infant visual sustained attention and respiratory sinus arrhythmia. *Child Development, 58*(2), 488-496.

Richards, J. E., & Casey, B. J. (1991). Heart rate variability during attention phases in young infants. *Psychophysiology, 28*(1), 43-53.

Rosenzweig, M. R., Leiman, A. L., & Breedlove, S. M. (1999). *Biological Psychology: An Introduction to Behavioral, Cognitive, and Clinical Neuroscience*. Sunderland, MA: Sinauer Associates.

Roth, J., Paris, S. G., & Turner, J. C. (2000). Students' perceived utility and reported use of test-taking strategies. *Issues in Education, 6*(1,2), 67-82.

Sanders, W. L., & Horn, S. P. (1998). Research findings from the Tennessee Value-Added Assessment System (TVAAS) database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education, 12*, 247-256.

Schore, A. N. (1994). *Affect Regulation and the Origin of the Self: The Neurobiology of Emotional Development*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Schroeder, L. (2004). *Student Drawing Assessment Protocol*. Unpublished document, Claremont Graduate University, Claremont, CA.

Schroeder, L. (2006). *What High-Stakes Testing Means for the Well-Being of Students and Teachers*. Doctoral dissertation, Faculty of Education, Claremont Graduate University, Claremont, CA.

Schroeder, L., Arguelles, L., & Bouman, S. (2006). Drawing to survive: A holistic assessment of student images of the high-stakes testing experience. Paper presented at the American Educational Research Association 2006 Annual Meeting, San Francisco, CA, April 8–12.

Sears, S. J., & Milburn, J. (1990). School-age stress. In L. E. Arnold (Ed.), *Childhood Stress* (pp. 224-226). New York: Wiley.

Spielberger, C. D. (1966). Theory and research on anxiety. In C. D. Spielberger (Ed.), *Anxiety and Behavior*. New York: Academic Press.

Spielberger, C. D. (1976). The nature and measurement of anxiety. In C. D. Spielberger & R. Diaz-Guerrero (Eds.), *Cross-Cultural Research on Anxiety*. Washington, D.C.: Hemisphere/Wiley.

Spielberger, C. D. (1980). *Test Anxiety Inventory*. Menlo Park, CA: Mind Garden.

Spielberger, C. D., & Vagg, P. R. (1995a). Test anxiety: A transactional process model. In C. D. Spielberger & P. R. Vagg (Eds.), *Test Anxiety: Theory, Assessment, and Treatment*. Washington, D.C.: Taylor & Francis.

Spielberger, C. D., & Vagg, P. R. (Eds.) (1995b). *Test Anxiety: Theory, Assessment, and Treatment*. Washington, D.C.: Taylor & Francis.

Stinchcombe, A. L. (1968). *Constructing Social Theories*. New York: Harcourt Brace & World.

Suess, P. E., Porges, S. W., & Plude, D. J. (1994). Cardiac vagal tone and sustained attention in school-age children. *Psychophysiology, 31*(1), 17-22.

Thayer, J. F., & Friedman, B. H. (1993). Assessment of anxiety using heart rate nonlinear dynamics. In W. Ditto (Ed.), *Chaos in Biology and Medicine: Proceedings of the International Society for Optical Engineering, 2036*, 42-48.

Thayer, J. F., & Friedman, B. H. (1997). The heart of anxiety: A dynamical systems approach. In A. Vingerhoets, F. van Bussel, & J. Boelhouwer (Eds.), *The (Non) Expression of Emotions in Health and Disease* (pp. 39-49). Tilburg, The Netherlands: Tilburg University Press.

Tiller, W. A., McCraty, R., & Atkinson, M. (1996). Cardiac coherence: A new, noninvasive measure of autonomic nervous system order. *Alternative Therapies in Health and Medicine, 2*(1), 52-65.

Turner, B. G., Beidel, D. C., Hughes, S., & Turner, M. W. (1993). Test anxiety in African American school children. *School Psychology Quarterly, 8*(2), 140-152.

Van Tilburg, E. (1987). Evaluation with a new twist. Using children's drawings to gather information. *Journal of Extension, 25*(4):5-7.

Wheelock, A., Bebell, D. J., & Haney, W. (2000a). Student self-portraits as test-takers: Variations, contextual differences, and assumptions about motivation. *Teachers College Record, November 2*: www.tcrecord.org, ID Number: 10635, Date Accessed: 10/19/2006.

Wheelock, A., Bebell, D. J., & Haney, W. (2000b). What can student drawings tell us about high-stakes testing in Massachusetts? *Teachers College Record, November 2*: www.tcrecord.org, ID Number: 10634, Date Accessed: 10/19/2006.

Wigfield, A., & Eccles, J. S. (1989). Test anxiety in elementary and secondary school students. *Educational Psychologist, 24*(2), 159-183.

Zeidner, M. (1998). *Test Anxiety: The State of the Art*. New York: Plenum Press.

Appendices

Appendix 1

Further Results from the Within-Groups Analysis in Chapter VII

Here we present the more detailed results of an ANCOVA in which test anxiety, test performance, and the SOS scales for each of the primary ethnic categories are broken down by Gender, Family Composition, and Class Academic Level for the experimental group and control group, respectively.

Experimental Group

Starting with the results presented in Table A.1.1, only one difference emerges by ethnic category for males and that is the significantly better CST ELA test performance of White Males over Hispanic Males (347.34 versus 333.73, $p < 0.05$). While a similar difference is also apparent for White Females over Hispanic Females (357.43 versus 346.50, $p < 0.01$), the latter report higher Life Preparedness than the former (3.17 versus 2.97, $p < 0.05$). However, Hispanic Females also report the lowest Extent of Friendship of the three ethnic groupings, which is significantly lower than that for the Other Ethnic group (3.26 versus 3.48, $p < 0.05$).

In terms of the breakdown of ethnicity by Family Composition, only two factors differentiate students in a Two (biological) Parent Family; White students outperformed Hispanic students on the CST ELA test (355.43 versus 345.12, $p < 0.05$), and White students from in-tact families had the most positive Feelings about School (3.66 versus 3.39 for intact-family students in the Other Ethnicity category; $p < 0.05$). By contrast more differences among the three ethnic categories are evident for students who live in an Other Parent Family situation. With the exception of the better CST ELA test performance of White students over Hispanics (346.20 versus 331.40, $p < 0.05$), the other three significant differences are between White students and those from an Other Ethnic group, and consistently the latter is higher than the former.

Thus the Other Ethnic group students from non-intact families are higher on Life Preparedness (3.31 versus 2.79, $p < 0.001$), on Positive Class Experience (3.16 versus 2.83, $p < 0.05$), and on Stress Management (2.52 versus 2.22, $p < 0.01$) than their counterparts in the Other Ethnic group.

Virtually the same factors are involved in differentiating the three ethnic groupings by for students in a Regular (non-advanced) Class. Again, on CST ELA test performance the White students score significantly higher than Hispanic students in a regular class (334.07 versus 323.49, $p < 0.01$). And while Hispanic students in a Regular Class were higher on Positive Class Experience (2.97 versus 2.79, $p < 0.05$) and Stress Management (2.45 versus 2.28, $p < 0.05$), their counterparts in an Other Ethnic Group were significantly higher than White students on Life Preparedness (3.20 versus 2.91, $p < 0.05$). By contrast only Life Preparedness differentiated the three ethnic groupings for those in an Advanced Class—Hispanic students were significantly higher than students in the Other Ethnic Group category (3.82 versus 2.83, $p < 0.01$). As we have done above, we treated the marginally significant difference on Educational Plans (data not shown) between Hispanic, White, and Other Ethnic Group students (3.91, 3.87, and 3.69, respectively; $p = 0.053$) as a chance result of the many statistical comparisons generated in the analysis.

Table A.1.1 Two-Way ANCOVA (Showing Significant Results) of Test Performance and SOS Scales by Selected Sociodemographic Variables for the Experimental Group

	1- Caucasian					2- Hispanic					3 -Other									Post Hoc Comparisons		
Dependent	N	Mean	SD	Adj. Mean	SEM	N	Mean	SD	Adj. Mean	SEM	N	Mean	SD	Adj. Mean	SEM	Mean Sq	F	$p <$	Sig.	1 vs 2	1 vs 3	2 vs 3
All Students																						
CST English - Language Arts 10	167	363.28	45.08	352.07	2.28	217	328.86	49.37	340.27	2.01	62	357.00	50.34	347.25	3.69	6172.25	7.37	0.001	0.001	0.001	ns	ns
Males																						
CST English - Language Arts 10	84	360.19	49.56	347.34	3.58	108	320.12	47.26	333.73	3.19	32	350.84	43.63	338.65	5.71	3955.52	3.85	0.05	0.023	0.05	ns	ns
Females																						
CST English - Language Arts 10	82	367.06	39.91	357.43	2.84	109	337.52	50.09	346.50	2.46	29	363.83	57.62	357.34	4.73	3118.27	4.83	0.01	0.009	0.05	ns	ns
Life Preparedness	87	2.94	0.73	2.97	0.06	115	3.24	0.63	3.17	0.05	29	2.82	0.87	3.02	0.10	1.05	3.62	0.05	0.028	0.05	ns	ns
Extent of Friendship	87	3.38	0.50	3.34	0.05	115	3.24	0.50	3.26	0.04	29	3.42	0.54	3.48	0.08	0.65	3.52	0.05	0.031	ns	ns	0.05
2 Biological Parent Families																						
CST English - Language Arts 10	107	368.45	46.17	355.43	2.83	142	332.28	48.40	345.12	2.47	35	365.77	55.92	353.45	4.87	3140.42	3.83	0.05	0.023	0.05	ns	ns
Feelings about School	111	3.67	0.75	3.66	0.05	155	3.58	0.76	3.57	0.05	36	3.31	0.62	3.39	0.10	0.98	3.05	0.05	0.049	ns	0.05	ns
Other Parent Families																						
CST English - Language Arts 10	60	354.07	41.88	346.20	3.84	75	322.39	50.85	331.40	3.46	27	345.63	40.23	338.06	5.70	3484.52	4.01	0.05	0.020	0.05	ns	ns
Life Preparedness	66	2.76	0.78	2.79	0.07	83	3.06	0.68	3.03	0.07	28	3.26	0.74	3.31	0.11	2.83	7.79	0.001	0.001	ns	0.001	ns
Positive Class Experience	63	2.86	0.77	2.83	0.07	80	2.93	0.61	2.96	0.06	25	3.19	0.77	3.16	0.11	0.99	3.36	0.05	0.037	ns	0.05	ns
Stress Management	67	2.20	0.68	2.22	0.07	82	2.48	0.63	2.47	0.06	28	2.51	0.50	2.52	0.10	1.46	4.93	0.01	0.008	0.05	0.05	ns
Regular Classes																						
CST English - Language Arts 10	108	342.41	37.52	334.07	2.61	185	317.34	41.19	323.49	2.00	45	335.49	34.37	330.20	4.00	3754.25	5.23	0.01	0.006	0.01	ns	ns
Life Preparedness	115	2.84	0.77	2.91	0.06	204	3.12	0.72	3.06	0.04	47	3.10	0.72	3.20	0.09	1.56	4.19	0.05	0.016	ns	0.05	ns
Positive Class Experience	110	2.83	0.78	2.79	0.06	194	2.94	0.71	2.97	0.05	42	2.98	0.78	2.97	0.10	1.22	3.17	0.05	0.043	0.05	ns	ns
Stress Management	117	2.31	0.70	2.28	0.05	203	2.44	0.67	2.45	0.04	47	2.44	0.64	2.47	0.08	1.18	3.86	0.05	0.022	0.05	ns	ns
Advancedd Classes																						
Life Preparedness	61	3.09	0.69	3.07	0.07	33	3.33	0.52	3.32	0.09	17	2.73	0.86	2.83	0.12	1.43	5.50	0.01	0.005	ns	ns	0.01

ANCOVA

The final step in this analysis involved comparing pre-post change differences on test anxiety, test performance, and SOS scales by Gender, Family Composition, and Academic Class Level categories for each of the three ethnic groupings. The results of the ANCOVA are presented in Table A.1.2.

Starting with the results for White students by gender, three factors are significant in differentiating Males from Females, all of which are higher for the female students: Extent of Friendship (3.13 versus 2.82, $p < 0.001$), Negative Affect (2.18 versus 1.96, $p < 0.01$), and Emotional Discord (2.19 versus 1.98, $p < 0.05$). For Hispanic students two different factors are significant, with Females reporting higher values than Males on Life Preparedness (3.23 versus 3.08, $p = 0.051$) and on Extent of Friendship (3.02 versus 2.80, $p < 0.01$). For students in the Other Ethnic Group Females there was a greater Extent of Friendship than that reported by Males (3.22 versus 2.91, $p < 0.05$). While this group of females achieved a marginally higher mean test score on CST ELA than Males (363.40 versus 351.23, $p = 0.06$), we will treat this result as likely due to chance.

Turning to Family Composition, while there are some differences for White and Other Ethnic Group students, no significant differences were found for the Hispanic students. For White students, significant differences favoring those from a Two (biological) Parent Family were found for Feelings about School (3.66 versus 3.45, $p < 0.05$) and Life Preparedness (3.02 versus 2.77, $p < 0.01$); the marginally significant difference observed for Stress Management (2.42 versus 2.25, $p = 0.06$) we will treat as a chance result. Interesting differences are also apparent for students in the Other Ethnic Group category, all of which favor those from Other Parent Family situations: Test Anxiety-Worry (1.78 versus 2.14, $p < 0.05$), Life Preparedness (3.23 versus 2.83, $p < 0.01$), Positive Class Experience (3.24 versus 2.85, $p < 0.05$), and Positive Affect (3.04 versus 2.74, $p < 0.05$).

Finally, moving to the results for Class Academic Level, apart from the expected higher CST ELA test performance of students in an Advanced Class across all three ethnic groupings, the strongest difference associated with Class Level was found for the White students. Of the three factors involved, Test Anxiety-Emotional was lower for those in Regular Classes (1.85 versus 2.08, $p < 0.05$), perhaps because they are not so concerned about their test scores. By contrast, Feelings about School and Positive Class Experience were found to be higher for those in an Advanced Class (3.73 versus 3.51, $p < 0.05$; and 3.12 versus 2.87, $p < 0.05$, respectively). For Hispanic students only, Life

Preparedness was associated with Class Level—it was higher for those in an Advanced Class (3.37 versus 3.12, p <0.05). Last, for students in the Other Ethnic Group category, only Test Anxiety-Emotional was associated with Class Level, with those in a Regular Class reporting lower test anxiety than their classmates in an Advanced Class (1.80 versus 2.22, p <0.05).

Table A.1.2 Two-Way ANCOVA (Showing Significant Results) of Test Anxiety, Test Performance and SOS Scales by Selected Sociodemographic Variables for the Experimental Group

Gender	Male					Female							
	N	Mean	SD	Adj. Mean	SEM	N	Mean	SD	Adj. Mean	SEM	Mean Sq	F	*p* <
White													
Extent of Friendship	89	2.58	0.68	2.82	0.06	87	3.38	0.50	3.13	0.06	2.73	11.25	0.001
Negative Affect	89	1.92	0.64	1.96	0.05	88	2.21	0.74	2.18	0.05	2.08	8.28	0.01
Emotional Discord	88	1.89	0.59	1.98	0.06	88	2.28	0.82	2.19	0.06	1.73	5.34	0.05
Hispanic													
Extent of Friendship	123	2.59	0.72	2.80	0.05	115	3.24	0.50	3.02	0.05	2.03	7.50	0.01
Other													
Extent of Friendship	34	2.74	0.76	2.91	0.08	29	3.42	0.54	3.22	0.09	1.36	6.61	0.05
Family Composition	2 Biological parents					Other parents							
	N	Mean	SD	Adj. Mean	SEM	N	Mean	SD	Adj. Mean	SEM	Mean Sq	F	*p* <
White													
Feelings about School	111	3.67	0.75	3.66	0.05	67	3.44	0.69	3.45	0.07	1.77	6.07	0.05
Life Preparedness	110	3.03	0.71	3.02	0.06	66	2.76	0.78	2.77	0.07	2.68	7.98	0.01
Other													
Test Anxiety-Worry	35	2.23	1.02	2.14	0.11	27	1.66	0.77	1.78	0.12	1.96	4.98	0.05
Life Preparedness	36	2.80	0.74	2.83	0.10	28	3.26	0.74	3.23	0.11	2.52	7.30	0.01
Positive Class Experience	34	2.88	0.81	2.85	0.10	25	3.19	0.77	3.24	0.12	2.18	6.38	0.05
Positive Affect	36	2.81	0.80	2.74	0.10	28	2.96	0.60	3.04	0.11	1.34	4.15	0.05
Class Level	Regular classes					Advanced Classes							
	N	Mean	SD	Adj. Mean	SEM	N	Mean	SD	Adj. Mean	SEM	Mean Sq	F	*p* <
White													
CST English - Language Arts 10	108	342.41	37.52	353.61	2.86	59	401.49	30.40	380.98	4.09	19361.06	25.69	0.001
Test Anxiety-Emotional	111	1.86	0.84	1.85	0.06	60	2.06	0.94	2.08	0.08	2.20	5.32	0.05
Feelings about School	117	3.44	0.72	3.51	0.05	61	3.86	0.69	3.73	0.07	1.84	6.32	0.05
Positive Class Experience	110	2.83	0.78	2.87	0.06	59	3.19	0.58	3.12	0.08	2.27	6.29	0.05
Hispanic													
CST English - Language Arts 10	185	317.34	41.19	324.75	2.16	32	395.47	39.22	352.63	5.79	15458.46	18.95	0.001
Life Preparedness	204	3.12	0.72	3.12	0.04	33	3.33	0.52	3.37	0.10	1.76	5.09	0.05
Other													
CST English - Language Arts 10	45	335.49	34.37	350.46	3.89	17	413.94	40.80	374.32	7.21	3892.09	6.98	0.05
Test Anxiety-Emotional	45	1.78	0.94	1.80	0.10	16	2.26	0.85	2.22	0.17	2.15	4.49	0.05
Educational Plans	47	3.72	0.48	3.75	0.05	17	3.63	0.75	3.53	0.09	0.59	4.42	0.05
Life Preparedness	47	3.10	0.72	3.10	0.09	17	2.73	0.86	2.71	0.14	1.93	5.44	0.05

ANCOVA

Control Group

The data presented in Table A.1.3 are discussed in Chapter VII, page 109.

Table A.1.3 Two-Way ANCOVA (Showing Significant Results) of Test Performance and SOS Scales by Selected Sociodemographic Variables for the Control Group

	White					Asian					Other							
Dependent	N	Mean	SD	Adj. Mean	SEM	N	Mean	SD	Adj. Mean	SEM	N	Mean	SD	Adj. Mean	SEM	Mean Sq	F	*p* <
All Students																		
Extent of Friendship	138	3.03	0.65	3.04	0.04	52	3.21	0.52	3.18	0.07	67	2.93	0.72	2.93	0.06	0.89	3.82	0.05
Interactional Difficulty	139	1.91	0.55	1.92	0.04	52	1.85	0.54	1.92	0.07	68	2.17	0.69	2.10	0.06	0.75	3.09	0.05
Males																		
Teacher Support	60	2.95	0.81	2.92	0.08	17	3.25	0.57	3.16	0.15	24	2.51	0.82	2.66	0.13	1.23	3.14	0.05
Extent of Friendship	60	2.69	0.65	2.69	0.07	17	2.96	0.53	2.99	0.13	23	2.48	0.72	2.46	0.12	1.33	4.40	0.05
2 Biological Parent Families																		
Extent of Friendship	85	3.04	0.67	3.09	0.05	40	3.23	0.51	3.19	0.07	42	2.95	0.70	2.87	0.07	1.15	5.85	0.01
Advanced Classes																		
CST English-Language Arts 10	72	418.29	43.21	413.87	3.22	42	401.81	34.23	405.21	4.20	24	390.29	34.30	397.61	5.57	2614.80	3.55	0.05
Extent of Friendship	72	3.11	0.66	3.11	0.05	43	3.24	0.50	3.21	0.06	25	2.83	0.66	2.90	0.09	0.78	4.36	0.05
Interactional Difficulty	73	1.78	0.48	1.78	0.05	43	1.85	0.55	1.86	0.06	25	2.06	0.60	2.03	0.09	0.56	3.16	0.05

ANCOVA

Moving to the data presented in Table A.1.4, we begin with the results of a within-category two-way ANCOVA in which the test anxiety, test performance, and SOS scales for each of the three primary ethnic categories (White, Asian, and Other Ethnicity) are broken down by Gender, Family Composition, and Class Academic Level (see Table A.1.4)

Of the few significant findings observed, Asian students consistently report a greater Extent of Friendship than the other ethnic categories. This is true for Asian Males (2.99 compared to 2.69 and 2.46 for Whites and Other Ethnic Group, respectively, p <0.05), for Asian student from intact families with both biological parents (3.19 compared to 3.09 and 2.87, respectively, p <0.01), and Asian students in an Advanced Class (3.21 compared to 3.11 and 2.90, respectively, p <0.05). Asian Males also have higher ratings for Teacher Support (3.16 compared to 2.92 and 2.66, respectively, p <0.05) and for Life Preparedness (3.23 compared to 2.95 and 2.77, p <0.05) than the other two ethnic groupings. And finally, while White students in an Advanced Class, on average, out perform their Asian and Other Ethnic Group classmates (413.87 versus 405.21 and 397.61, respectively, p <0.05), they also report the lowest level of Interactional Difficulty (1.78 compared to 1.86 for Asian and 2.03 for Other Ethic Group, p <0.05).

Table A.1.4 Two-Way ANCOVA (Showing Significant Results) of Test Performance and SOS Scales by Selected Sociodemographic Variables for the Control Group

	White					Asian					Other							
	N	Mean	SD	Adj. Mean	SEM	N	Mean	SD	Adj. Mean	SEM	N	Mean	SD	Adj. Mean	SEM	Mean Sq	F	p
All Students																		
Extent of Friendship	138	3.03	0.65	3.04	0.04	52	3.21	0.52	3.18	0.07	67	2.93	0.72	2.93	0.06	0.89	3.82	0.
Interactional Difficulty	139	1.91	0.55	1.92	0.04	52	1.85	0.54	1.92	0.07	68	2.17	0.69	2.10	0.06	0.75	3.09	0.
Males																		
Teacher Support	60	2.95	0.81	2.92	0.08	17	3.25	0.57	3.16	0.15	24	2.51	0.82	2.66	0.13	1.23	3.14	0.
Extent of Friendship	60	2.69	0.65	2.69	0.07	17	2.96	0.53	2.99	0.13	23	2.48	0.72	2.46	0.12	1.33	4.40	0.
2 Biological Parent Families																		
Extent of Friendship	85	3.04	0.67	3.09	0.05	40	3.23	0.51	3.19	0.07	42	2.95	0.70	2.87	0.07	1.15	5.85	0.
Advanced Classes																		
CST English-Language Arts 10	72	418.29	43.21	413.87	3.22	42	401.81	34.23	405.21	4.20	24	390.29	34.30	397.61	5.57	2614.80	3.55	0.
Extent of Friendship	72	3.11	0.68	3.11	0.05	43	3.24	0.50	3.21	0.06	25	2.83	0.68	2.90	0.09	0.78	4.36	0.
Interactional Difficulty	73	1.78	0.48	1.78	0.05	43	1.85	0.55	1.86	0.06	25	2.06	0.60	2.03	0.09	0.56	3.16	0.

The final step involved comparing pre-post change differences on test anxiety, test performance, and SOS scales by Gender, Family Composition, and Academic Class Level categories for each of the three ethnic groupings. The results of the ANCOVA are presented in Table A.1.5.

Table A.1.5 Two-Way ANCOVA (Showing Significant Results) of Test Anxiety, Test Performance and SOS Scales by Selected Sociodemographic Variables for the Control Group

Gender	Male					Female							
	N	Mean	SD	Adj. Mean	SEM	N	Mean	SD	Adj. Mean	SEM	Mean Sq	F	p <
White													
Test Anxiety-Global	58	1.82	0.83	2.03	0.08	77	2.40	0.84	2.24	0.07	1.42	4.40	0.05
Test Anxiety-Worry	58	1.77	0.83	1.91	0.09	77	2.32	0.89	2.22	0.07	2.90	7.08	0.01
Extent of Friendship	60	2.69	0.65	2.91	0.06	77	3.29	0.52	3.12	0.05	1.09	5.44	0.05
Asian													
Parental Support	16	3.72	0.38	3.74	0.09	34	3.50	0.60	3.49	0.06	0.70	5.17	0.05
Other													
Test Anxiety-Worry	24	1.97	0.75	2.18	0.14	39	2.73	0.93	2.60	0.11	2.39	5.72	0.05
Teacher Support	24	2.51	0.82	2.61	0.14	41	3.10	0.87	3.04	0.11	2.82	6.01	0.05
Extent of Friendship	23	2.48	0.72	2.65	0.14	41	3.14	0.61	3.05	0.10	1.95	5.28	0.05
Class Level	Regular classes					Advanced Classes							
	N	Mean	SD	Adj. Mean	SEM	N	Mean	SD	Adj. Mean	SEM	Mean Sq	F	p <
White													
CST English - Language Arts 10	61	351.39	50.54	376.46	4.01	72	418.29	43.21	397.05	3.64	10055.42	12.43	0.001
Teacher Support	66	2.86	0.84	2.92	0.07	73	3.22	0.71	3.16	0.07	1.98	5.83	0.05
Educational Plans	65	3.62	0.57	3.68	0.04	73	3.89	0.26	3.84	0.04	0.84	7.82	0.01
Interactional Difficulty	66	2.06	0.59	2.00	0.06	73	1.78	0.48	1.83	0.05	0.89	4.61	0.05

ANCOVA

Starting with the results for White students by gender, Females are higher for test anxiety (Test Anxiety-Global: 2.24 versus 2.03, $p < 0.05$) and have a greater Extent of Friendship (3.12 versus 2.91, $p < 0.05$) than Males. For Asian students, Males are significantly higher on Parental Support (3.74 versus 3.49, $p < 0.05$) than their female counterparts; we regard the marginally significant result on Life Preparedness (3.24 versus 2.94, $p = 0.052$) as a chance result. And for those in the Other Ethnic Group category, Females have higher Test Anxiety-Worry (2.60 versus 2.18, $p < 0.05$), higher Teacher Support (3.04 versus 2.61, $p < 0.05$), and a greater Extent of Friendship (3.05 versus 2.65, $p < 0.05$) than their Male counterparts. And while no differences were found when the three ethnic categories were broken down by Family Composition, the only significant finding involving Class Level was for the White students. For this ethnic category those in an Advanced Class had a higher mean score on the CST ELA (397.05 versus 376.46, $p < 0.001$), reported more Teacher Support (3.16 versus 2.92, $p < 0.05$), more optimistic Educational Plans (3.84 versus 3.68, $p < 0.01$), and reported less Interactional Difficulty (1.83 versus 2.00, $p < 0.05$) than their counterparts in a Regular Class by the end of the study.

Appendix 2

Table A.2.1 Within-Groups Analysis: Summary of Significant Results by Intervention Status

Pre – Post	CONTROL	EXPERIMENTAL
CST – ELA 10	--	Post < Pre*
TAI – Relationship	Post < Pre**	Post < Pre***
TAI – Worry	Post < Pre*	Post < Pre***
TAI – Emotional	Post < Pre*	Post < Pre***
Educational Plans	--	Post < Pre[m]
Life Preparedness	--	Post < Pre*
Positive Affect	--	Post > Pre*
Stress Management	Post > Pre**	--
GENDER		
TAI – Global	Female > Male**	--
TAI – Worry	Female > Male***	--
Extent of Friendship	Female > Male**	Female > Male***
Negative Affect	--	Female > Male***
Emotional Discord	Female > Male*	Female > Male*
Family Composition		
TAI – Global	--	Other Parent < 2 Parents*
TAI – Worry	--	Other Parent < 2 Parents*
Teacher Support	--	Other Parent < 2 Parents[m]
Ethnicity		
CST – ELA 10	--	White > Other > Hispanic***
Extent of Friendship	Asian > White > Other*	--
Interactional Difficulty	Asian < White = Other*	--
Class Level		
CST – ELA 10	Advanced > Regular***	Advanced > Regular****
Feelings About School	--	Advanced > Regular*
Teacher Support	Advanced > Regular**	--
Educational Plans	Advanced > Regular***	--
Interactional Difficulty	Advanced < Regular[m]	--

[m] $p < 0.1$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$

Table A.2.1 continued…

	CONTROL	EXPERIMENTAL
MALE:		
CST – ELA 10	--	White > Other > Hispanic***
Teacher Support	Asian > White > Other*	--
Life Preparedness	Asian > White > Other[m]	--
Extent of Friendship	Asian > White > Other*	--
Female:		
CST – ELA 10	--	White ≅ Other > Hispanic**
Life Preparedness	--	Hispanic > Other > White*
Extent of Friendship	--	Other > White > Hispanic*
2-Parent Family:		
CST – ELA 10	--	White > Other > Hispanic*
Feelings About School	--	White > Hispanic > Other*
Extent of Friendship	Asian > White > Other**	--
Other Parent Family:		
CST – ELA 10	--	White > Other > Hispanic*
Life Preparedness	--	Other > Hispanic > White***
Pos. Class Experience	--	Other > Hispanic > White*
Stress Management	--	Other > Hispanic > White**
Regular Class:		
CST – ELA 10	--	White > Other > Hispanic**
TAI – Emotional	--	Other < White < Hispanic*
Life Preparedness	--	Other > Hispanic > White*
Pos. Class Experience	--	Hispanic = Other > White*
Stress Management	--	Other > Hispanic > White*
Advanced Class:		
CST – ELA 10	White > Asian > Other*	--
Educational Plans	--	Hispanic > White >Other[m]
Life Preparedness	--	Hispanic > White >Other**
Extent of Friendship	Asian > White > Other*	--
Interactional Difficulty	White < Asian < Other*	--

[m] $p < 0.1$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$

Table A.2.1 continued…

	CONTROL	Experimental
White (by Gender):		
TAI – Global	Female > Male*	--
TAI – Worry	Female > Male**	--
Extent of Friendship	Female > Male*	Female > Male***
Negative Affect	--	Female > Male**
Emotional Discord	--	Female > Male*
Asian (by Gender):		**Hispanic (by Gender):**
Life Preparedness	Male > Female^{m}	Female > Male^{m}
Parental Support	Male > Female*	Female > Male**
Extent of Friendship	--	--
Other Ethnicity (by Gender):		
CST – ELA 10	--	Female > Male^{m}
TAI – Worry	Female > Male*	--
Teacher Support	Female > Male*	--
Extent of Friendship	Female > Male*	Female > Male*
White (by Family. Composition):		
Feelings About School	--	2 Parents > Other Parent*
Life Preparedness	--	2 Parents > Other Parent**
Stress Management	--	2 Parents > Other Parent^{m}
Asian (by Family Comp.):		**Hispanic (by Family Comp.):**
	No Significant Findings	No Significant Findings
Other Ethnic. (by Family Comp.):		
TAI – Worry	--	2 Parent > Other Parent*
Life Preparedness	--	2 Parent < Other Parent**
Positive Class Experience	--	2 Parent < Other Parent*
Positive Affect	--	2 Parent < Other Parent*
White (by Class Level):		
CST – ELA 10	Advanced > Regular***	Advanced > Regular***
TAI – Emotional	--	Regular < Advanced*
Feelings About School	--	Advanced > Regular*
Teacher Support	Advanced > Regular*	Advanced > Regular*
Educational Plans	Advanced > Regular**	--
Positive Class Experience	--	Advanced > Regular*
Interactional Difficulty	Regular > Advanced*	--
Asian (by Class Level):		**Hispanic (by Class Level):**
CST – ELA 10	Advanced > Regular^{m}	Advanced > Regular***
Life Preparedness	--	Advanced > Regular*
Other Ethnic. (by Class Level):		
CST – ELA 10	--	Advanced > Regular*
TAI – Emotional	--	Regular < Advanced*

^{m} $p < 0.1$, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, $****p < 0.0001$

Appendix 3

Student Opinion Survey

Dear Student,

You are being asked to complete this survey as part of an important study funded by the U.S. Department of Education to evaluate the implementation of the TestEdge program in schools across the country. Your parents have given permission for you to participate in this study. While we hope that you will choose to participate, you still get to decide if you want to be part of this research study. You can choose not to participate or stop at anytime if you change your mind later.

If you complete this survey, no one at your school will see the answers you give. Your answers are anonymous and confidential. **Do not put your name on the survey form**. You can also skip any of the questions you do not choose to answer.

When you have completed the survey, hand it back to the research assistant who gave you the form. This will signify that you have freely given us your permission to be included in the study.

Sincerely,

The Research Team

INSTITUTE OF HEARTMATH®
Empowering Heart-Based Living

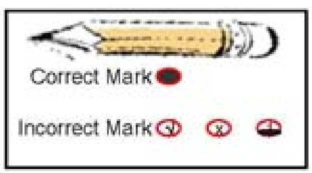
DIRECTIONS:

This survey asks your opinion about this school and your life over the past few weeks. Each question is asked for a specific reason, and we would appreciate it if you would answer all questions; your opinion matters.

To ensure confidentiality, please do not write your name on this form. Do not identify yourself in any other way.

Please fill in completely the oval for the answer that best reflects your opinion. Mark only one answer for each question.

There are no right or wrong answers; this is not a test. Answer each question the way that best describes you or your situation. While you do not have to answer all the questions, we hope that you will because your opinion is important.



Use a **Number 2 pencil** only.
Make heavy black marks, **fill the oval completely.**
Check to be sure you have answered every question.
Do not make stray marks on the questionnaire.
Do not fold or wrinkle your questionnaire.

Thank you!

8 Test Anxiety Inventory questions included by permission ©1980 Charles D. Spielberger Published by Mind Garden
12 Questions from California Health Kids Survey included by permission ©2003 CA Dept. of Ed.
Student Opinion Survey form ©2005 Institute of HeartMath

1

Section 1. Background Information

The following questions ask for some background information about you.

Today's date:

Month	Day	Year

1. How old are you?

- ① 10 years old or younger
- ② 11 years old
- ③ 12 years old
- ④ 13 years old
- ⑤ 14 years old
- ⑥ 15 years old
- ⑦ 16 years old
- ⑧ 17 years old
- ⑨ 18 years old or older

2. What is your gender?

- ① Male
- ② Female

3. Mark **one** of the following that best describes you.

- ① American Indian or Alaska Native
- ② Native Hawaiian or Pacific Islander
- ③ Asian
- ④ Black or African American
- ⑤ Hispanic or Latino (includes Cuban, Mexican, Chicano, Puerto Rican, etc.)
- ⑥ White or Caucasian (non-Hispanic)
- ⑦ Other_______________________________

4. Which of the following best describes your present family situation at home?

I live:

- ① in a two-parent family with both of my biological parents
- ② in a mixed family with one of my parents
- ③ in a single-parent family with one of my parents
- ④ in dual custody: I live with each of my parents at different times
- ⑤ in a family with adult relatives, but not my parents
- ⑥ in a family with adults who are not related to me
- ⑦ in a family with no adults
- ⑧ I live alone
- ⑨ Other _______________________________

5. How long have you been at this school?

- ① This is my 1st year
- ② This is my 2nd year
- ③ This is my 3rd year
- ④ This is my 4th year

2

Section 2. Opinions About School and Home

Answer the following questions about your SCHOOL.

	Strongly Disagree	Disagree	Neither Disagree Nor Agree	Agree	Strongly Agree
6. I feel close to people at this school	1	2	3	4	5
7. I am happy to be at this school	1	2	3	4	5
8. I feel safe in this school	1	2	3	4	5

At my school, there is a teacher or some other adult…

	Not at All True	A Little True	Pretty Much True	Very Much True
9. who really cares about me	1	2	3	4
10. who listens to me when I have something to say	1	2	3	4
11. who believes that I will be a success	1	2	3	4

How true do you feel these statements are about you personally?

	Not at All True	A Little True	Pretty Much True	Very Much True
12. I plan to graduate from high school	1	2	3	4
13. I plan to go to college or some other school after high school	1	2	3	4
14. There is a purpose to my life	1	2	3	4

I feel what I have learned at this SCHOOL:

	Not at All True	A Little True	Pretty Much True	Very Much True
15. will help me be successful as an adult	1	2	3	4
16. has prepared me to do well at college or get a good job	1	2	3	4
17. has inspired me to want more for myself out of life	1	2	3	4

How true are these statements about your HOME?

In my home, there is a parent or some other adult…

	Not at All True	A Little True	Pretty Much True	Very Much True
18. who is interested in my schoolwork	1	2	3	4
19. who believes that I will be a success	1	2	3	4
20. who talks with me about my problems	1	2	3	4
21. who loves me no matter what	1	2	3	4

Section 3. Feelings About Your Class, Classmates, and Friends

Read the statements below and indicate how you feel while you are <u>in this class</u>.

	Not at All True	A Little True	Pretty Much True	Very Much True
22. I enjoy this class and find it fun	1	2	3	4
23. I am pleased with how much I am learning	1	2	3	4
24. I feel that there are mostly good feelings among all of us in this class	1	2	3	4
25. I feel the teacher cares about me and my classmates as individuals	1	2	3	4

3

Take a moment to think about your relationships with your close friends in three different social settings: in this class, in this school, and in your community.

Indicate how many close friends you have in these different social settings.

26. In this **class** *(Please indicate how many)*

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
>30

27. In the rest of this **school** (outside this class) *(Please indicate how many)*

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
>30

28. Outside this school, in the **community** *(Please indicate how many)*

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
>30

For those classmates or friends you feel closest to, do you:

	Almost Never	Sometimes	Often	Almost Always
29. Do homework together	1	2	3	4
30. Feel very comfortable with each other	1	2	3	4
31. Talk on the phone	1	2	3	4
32. Confide in each other about personal situations or problems	1	2	3	4
33. Spend time alone together as friends	1	2	3	4
34. Have strong feelings about each other	1	2	3	4
35. Care a lot about each other	1	2	3	4

Section 4. Feelings About Your Life

In thinking about my life over the past few weeks, I generally find that:

	Almost Never	Sometimes	Often	Almost Always
36. I feel happy	1	2	3	4
37. I feel stressed	1	2	3	4
38. I feel lonely	1	2	3	4
39. I feel calm	1	2	3	4
40. I feel anxious	1	2	3	4
41. I feel appreciative	1	2	3	4
42. I feel cheerful	1	2	3	4
43. I feel sad	1	2	3	4
44. I feel angry	1	2	3	4
45. I feel enthusiastic	1	2	3	4
46. I feel hopeful	1	2	3	4
47. I feel depressed	1	2	3	4
48. I feel loved	1	2	3	4
49. I feel disappointed	1	2	3	4

In thinking about my feelings over the past few weeks, I generally find that:

50. I feel powerless over what I am feeling	1	2	3	4
51. I don't always know clearly how I feel	1	2	3	4
52. I have opposite feelings from one moment to the next	1	2	3	4
53. I keep negative feelings bottled up inside	1	2	3	4
54. I feel overwhelmed by my feelings	1	2	3	4

4

In my interactions with others, there are situations in which:

	Almost Never	Sometimes	Often	Almost Always
55. I find it difficult to know what others are feeling	1	2	3	4
56. I have difficulty sharing my feelings with others	1	2	3	4
57. I don't feel that I am being heard and understood	1	2	3	4
58. I feel that I don't matter	1	2	3	4
59. I get into arguments or fights	1	2	3	4

When I am stressed out, I:

	Almost Never	Sometimes	Often	Almost Always
60. Try to manage my stress in the moment it happens	1	2	3	4
61. Talk to a friend, family member, or teacher	1	2	3	4
62. Watch TV or get on the Internet to distract myself from my feelings	1	2	3	4
63. Try to breathe deeply or relax to calm myself down	1	2	3	4
64. Find I want to sleep more than usual	1	2	3	4
65. Find it helpful to do physical exercise or some other activity	1	2	3	4

To handle my stress:

	Almost Never	Sometimes	Often	Almost Always
66. I try to change to positive feelings when I've had negative feelings for a while	1	2	3	4
67. I am able to calm myself down when feeling distressed	1	2	3	4
68. I am able to control my feelings effectively	1	2	3	4
69. I try to be less judgmental when feeling angry at myself	1	2	3	4
70. I try to change to hopeful feelings when feeling depressed	1	2	3	4
71. I try to become more peaceful when I'm feeling anxious	1	2	3	4

Section 5. Feelings About Tests

Following are a number of statements that people have used to describe their feelings about tests.

Indicate how you generally feel:

	Almost Never	Sometimes	Often	Almost Always
72. I freeze up on important exams	1	2	3	4
73. The harder I work at taking a test, the more confused I get	1	2	3	4
74. During tests I find myself thinking about consequences of failing	1	2	3	4
75. During examinations I get so nervous that I forget facts I really know	1	2	3	4
76. I feel very jittery when taking an important test	1	2	3	4
77. During tests I feel very tense	1	2	3	4
78. I feel very panicky when I take an important test	1	2	3	4
79. I worry a great deal before taking an important examination	1	2	3	4

5

80. **Please take a few minutes to draw or sketch a picture that shows how you feel when taking an important test. (Use the space provided below.)**

Thank you for taking this survey!

6


INSTITUTE OF HEARTMATH®
Empowering Heart-Based Living
14700 West Park Avenue, Boulder Creek, California
831-338-8500 • www.heartmath.org