Toward a Deeper Understanding of Student Performance in Virtual High School Courses:
Using Quantitative Analyses and Data Visualization to Inform Decision Making

W. Patrick Dickson, Ph.D.
Principal Investigator

Report Submitted to
Michigan Virtual University

http://www.mivu.org

July 21, 2005

This study was funded by a grant awarded in 2004 from the North Central Regional Laboratory (NCREL), 1120 East Diehl Road, Suite 200, Naperville, Illinois 60563. The content of this report does not necessarily reflect the position or policy of NCREL or the U.S. Department of Education, nor does mention of trade names, commercial products or organizations imply endorsement by NCREL or the federal government.
July 18, 2005

Mr. Andrew Henry
Chief Officer, Research and Development Group
Learning Point Associates
1120 E. Diehl Rd. Suite 200
Naperville, IL 60563

Dear Mr. Henry:

In December 2004 the Michigan Virtual University (MVU) was awarded a K-12 Online Learning Research Initiative Grant from the North Central Regional Educational Laboratory (NCREL). I am pleased to transmit the final report of this innovative study, *Toward a Deeper Understanding of Student Performance in Virtual High School Courses: Using Quantitative Analyses and Data Visualization to Inform Decision Making*, for your review.

The Michigan Legislature established the Michigan Virtual High School (MVHS) in 2000 and over the past five years the MVHS has grown to become a viable educational option for students and schools across the state. As a result of the NCREL grant, the MVU has been able to engage external and independent expertise to analyze existing MVHS demographic and enrollment data, and to develop strategies that allow greater access to data for decision making on courses, instructors and scheduling. Dr. Patrick Dickson, a professor in the College of Education at Michigan State University served as the principal investigator and led this study under a contract with MVU. Dr. Dickson worked closely with MVU’s leadership team and the MVHS staff to develop the findings and recommendations included in this report. This research has far-reaching implications for the MVHS and has already produced a number of internal quality improvements related to our use of data.

In these times of budgetary constraints and rising demands for accountability, state policy makers, legislative leaders and educators across the U.S. are seeking information on the impact of investments in online learning. The NCREL grant has served as a catalyst for the MVU to analyze an extensive amount of information on the growth and development of the MVHS, and on the performance of the students served. This pioneering work is an essential step in building a continuous improvement model for e-learning organizations. The principles and techniques for data visualization illustrated in this report have broader applicability to traditional schools as well.

We are grateful to NCREL for the support provided over the past year and we look forward to future opportunities to collaborate with NCREL and Learning Point Associates. Please let me know if we can provide any additional information.

Sincerely,

Jamey Fitzpatrick
Interim President
Acknowledgements

This report could not have been written without extensive contributions from many people. They deserve credit for what is good about the report. The author accepts responsibility for the remainder.

Jamey Fitzpatrick, Interim President of MVU provided leadership that created a climate of openness to the explorations of data and provided extensive feedback on the report itself. Dan Schultz contributed valuable ideas and background information that helped contextualize and inform the analyses. He also provided close reading of the many drafts. Bob Currie clarified aspects of the organization of MVHS. Melanie Rodgers worked many hours to organize the data for the analyses. Brandon Bautista helped in the development of Excel data files. Jason Marentette provided detailed analyses in Access as input to Excel. Justin Meese developed data visualizations in MapPoint. Kirby Milton facilitated access to MVU resources. Londa Horton helped interpret information about the teaching staff. The entire MVU staff provided clerical support and congenial company.

Andrew Dickson quickly mastered xCelsius to develop a prototype for interactive dashboard elements. Liz Weber shared her expertise with arcane features of Excel. Debra Lashbrook provided invaluable contributions to the layout and final revisions of the report.

We would also like to acknowledge and thank our colleagues at NCREL for their encouragement. Andrew Henry provided the initial leadership that led to this project. Bob Blomeyer has been tireless in his support for everyone involved in this project. John Watson, Rosina Smith and Tom Clark have been infinitely patient and thoughtful.
## Table of Contents

Cover Page ........................................................................................................................................... 1  
Letter of Transmittal from MVU to NCREL ......................................................................................... 2  
Acknowledgements ........................................................................................................................ 3  
Table of Contents ............................................................................................................................. 4  
List of Figures ...................................................................................................................................... 6  

Abstract ............................................................................................................................................. 7  
Executive Summary ........................................................................................................................ 8  

Introduction ...................................................................................................................................... 13  
  A Look at Michigan Virtual High School .................................................................................... 13  
  Context of the MVHS Study ........................................................................................................ 14  
  Meta-Analyses of Online Versus Traditional Courses ............................................................ 14  
  Statistical and Practical Implications of Large Variability ....................................................... 15  
  The Present Study ....................................................................................................................... 16  

Conceptual Background for the Analysis ......................................................................................... 16  
  Overview of Analysis .................................................................................................................. 16  
  Visual Display of Quantitative Data .......................................................................................... 16  
  Data-Driven Decision Making in Real-time: Potential for Virtual Courses ............................ 17  
  Examples of Other Work on Data-Driven Decision Making .................................................. 18  
  Data Dashboards and the “Balanced Scorecard” ....................................................................... 18  
  Sentinel Events: An Example from the Medical Profession .................................................... 19  
  Creating a Climate for Data-Driven Decision Making in K-12 Education .............................. 19  
  Major Corporate Providers of Data on K-12 Schools ............................................................ 19  
  The Center for Educational Performance and Information .................................................. 20  
  Home-Grown Solutions or Commercial Packages for D3? ...................................................... 20  
  Software for Data Visualization: From Access to Excel ......................................................... 20  
  Summary ....................................................................................................................................... 22  

Analysis of Enrollment and Student Performance Data .................................................................. 23  
  Visualizing the Big Picture: Placing the Work of MVHS in Geographic Context .................. 23  
  Changes over Time: Placing MVHS in Temporal Context ....................................................... 25  
    Enrollment Trends over Five Years: The “Big Picture” ......................................................... 26  
    Virtual Summer, Virtual School ............................................................................................. 27  
  Districts and Schools Served as Level of Analysis .................................................................... 28  
  Beyond Time and Space: Toward a Fine-Grained Look at Recent Semester .......................... 29  
  Focusing in on MVHS Courses: Spring 2004 through Spring 2005 ...................................... 30  
    Courses by Major Subject Area: The “Big Picture” of MVHS’ Curriculum ....................... 30  
    Understanding Sources of Variability in Online Courses .................................................... 31  
  Student Performance in MVHS Courses: “Final Score” as Percent ....................................... 31  
  Average Final Score by Major Subject Matter Category ......................................................... 31  
  Variation within World Languages by Language ....................................................................... 33
Variability Among Science Courses: Two Distinct Sources of Variability ..........35
What Do the Bar Charts Fail to Tell You? .................................................................38
Comparing Two Distributions in One Bar Chart.........................................................39
Pushing Microsoft Excel to Display Distributions ...................................................40
Displaying a Distribution with Excel by Plotting Against a Constant.........................41

Beyond Final Scores: Using Behavioral Data in Online Courses............................42
Students in Online Courses Generate Data in Real-Time......................................42
Examples of Poor Use of Data....................................................................................42
Blackboard's Pie Charts: Mixing Instructor Data with Student Data.........................42
Poor Use of Color and Poor Organization of Categories........................................43
Failure to Do the Work of Data Analysis for Instructors........................................44
Predicting Success in Online Courses from Active Participation (“Clicks”)..............44
Teachers Teach Students, Not Dots..........................................................................46
Use the Computer to Inform Students Directly.......................................................46
When Are Students Online in MVHS Courses?.......................................................47
Access by Hour of Day: Two Examples of Variability.............................................47
Access by Day of Week: Two Examples of Variability............................................49
Designing Displays of Course Statistics to Speak Directly to Students.....................51
Data Displays Showing Individual Students: Personal Learning Trajectories..........52

Toward Real-Time Data to Inform Decision Making................................................53
Data Dashboards .......................................................................................................53
Supporting Online Teachers' Learning: Data Design for Pedagogical Reflection ...54

Conclusions..................................................................................................................55
Designing Graphics to Promote Thought and Discussion ........................................56
Return on Investment: Making Data Useful at Its Point of Origin...........................56
Effective Models for Valid and Fruitful Scientific Research on Online Learning ....57
Design Experiments in Online Courses: Continuous Quality Improvement...........57
Intractable Contradictions at Heart of Meta-Analyses & Other Data Aggregation...58
A Partial Solution to this Intractable Dilemma.........................................................59
Opportunities for Further Research .........................................................................60

References..................................................................................................................61

Appendix A: Biographical Information .....................................................................63
Appendix B: MVHS: A Data Dashboard Prototype in Excel....................................64
List of Figures

Figure 1. Distribution of MVHS Students Across the State and Beyond ...............23
Figure 2. Location of MVHS Students in Grand Rapids Area ............................. 24
Figure 3. MVHS Enrollments from Fall 2000 through Summer 2005 .................. 26
Figure 4. Enrollment in MVHS Summer School versus School Year Courses ...... 27
Figure 5. Change in Districts and Schools Served by Year ............................... 28
Figure 6. Sections of Courses Offered by MVHS by Major Subject Matter Area .. 30
Figure 7. Average Final Score by Subject Matter Area ..................................... 32
Figure 8. Average Final Score in World Language Courses by Language ............ 34
Figure 9. Comparison of Final Scores by First vs Second Course in Sequence ...... 35
Figure 10. Distribution of Individual Scores as Bar Chart ................................. 38
Figure 11. Distribution of Scores in First (A) vs Second (B) Course in Sequence 40
Figure 12. Distribution of Final Scores in Three Sections of One Course .......... 41
Figure 13. Pie Chart Showing Total Clicks in One Blackboard Course ............... 43
Figure 14. Scatterplot of Relationship between Total Clicks and Final Score ...... 45
Figure 15a. Access by Hour of Day in One MVHS Course ............................... 47
Figure 15b. Access by Time of Day in Another MVHS Course .......................... 48
Figure 16a: Access by Day of Week: Course A ............................................. 49
Figure 16b: Access by Day of Week: Course B ............................................. 50
Figure 17. Trends for Individuals by Name by Month ..................................... 52
Toward a Deeper Understanding of Student Performance in Virtual High School Courses: Using Quantitative Analyses and Data Visualization to Inform Decision Making

Abstract

During the spring of 2005 the Michigan Virtual University (MVU) received a grant from the North Central Regional Educational Laboratory (NCREL) to explore ways in which data analysis and visualization could be used to better understand student performance in online learning environments. This grant was one of seven studies funded to study online learning in virtual high schools around the country. The purpose of the Michigan Virtual High School (MVHS) study was to determine whether a fine grained look at data would help shed light on the large degree of variability of student performance in virtual courses. This variability is noted frequently in meta-analyses comparing online and face-to-face classes. The project involved close collaboration with MVHS staff to organize existing data from courses over the past five years into a single database. A major objective of this project was to develop new ways of looking at data, with a view to informing decision making. The purpose was not to evaluate MVHS or specific courses.

Analyses used visual displays of quantitative data displays to shed light on variability in student performance. Initial analyses of MVHS trends in enrollments over a five-year period by semester and location helped contextualize the next phase of analysis. In this phase detailed analyses were carried out on average student performance by subject matter areas (such as mathematics, science and world languages). Analyses within these subject matter areas analyzed in even greater detail the distribution of individual students' final scores within specific courses. These analyses revealed that the variation in students' final scores varied considerably across the types of courses. In some cases the distributions were bimodal with many students doing well and some doing poorly or dropping out. Bimodal distributions violate assumptions of meta-analysis.

An iterative process of MVHS data analysis, data visualization and discussion with staff clarified the need for real-time data displays that can be readily and effectively used by educators and decision makers. These analyses offered the beginning of an understanding of the question, What data, what decisions? For example, does the data display offer an appropriate level of detail in the data for a specific stakeholder? For a teacher, a display of the grade distribution of 20 current students is more immediately engaging than a bar graph showing the mean of all 4th grade scores on a statewide test taken by students one or two years ago. Two data design principles influence impact. Displays should include a meaningful contrast and compare results with those of an affectively significant reference group. Examples of data displays that evoke reflection and raise different kinds of questions are presented.

The report places the MVHS study in the context of the burgeoning demands for accountability and data-driven decision making. The report suggests that K-12 educators have much to learn from sophisticated data tools now used in the business world, such as data dashboards, balanced scorecards, and sentinel events. Of course, these tools need to be adapted to fit the needs and context of K-12 schools. The report concludes with a discussion of how this project contributed to strengthening the culture of data at MVHS and offers some recommendations for how future work in this area could strengthen the usefulness of data at the school and district levels and thus also contribute to creating higher quality data for other types of analyses at higher levels.
Executive Summary

During the spring of 2005 the Michigan Virtual University (MVU) received a grant from the North Central Regional Educational Laboratory (NCREL) to carry out an intensive analysis of available data on performance of students. The purpose of this investigation was to determine whether a fine-grained look at this data would help shed light on the reports of no significant difference between online and traditional courses in recent meta-analyses.

The first phase of the study focused on working with the Michigan Virtual High School (MVHS) staff to organize the available data in preparation for analysis. One outcome of this phase was the development of a better understanding of the types and quality of existing MVHS data, which in turn contributed to procedures for routinely organizing such data in the future.

While working to organize existing data, we began exploring literature on uses of data in other areas, especially in business and science. In the business world there is extensive and growing literature on “data-driven decision making,” which is receiving increased attention in education. In the scientific realm new tools for data visualization have become essential tools for scientists seeking to understand complex phenomena, where work by Edward Tufte (2001) on the visual display of quantitative data has been influential (http://www.edwardtufte.com).

We began seeking ways of applying ideas from the worlds of business and science to our intensive analysis of data on student performance in online courses at MVHS, with a focus on the mirror-image questions of What data, what decisions? and What decisions, what data? The first implies what decisions might be better informed when working with existing data. The second question asks what kinds of decisions are available to administrators, teachers and students that might be better informed if certain types of data were readily accessible. Throughout this process, we continued to emphasize the importance of asking how data might be used to inform the stakeholders in any virtual high school.

The outcomes of these investigations led to several useful accomplishments for the MVHS specifically, and highlighted some potentially important conclusions for the broader research community. First, organizing and checking the accuracy of data on student performance required considerable effort on the part of MVHS staff. Like many startup organizations, the primary challenge for MVHS during this period, by necessity, was on launching the organization. During this time there were also changes in personnel and in the computer systems used for enrolling and recording student scores. These changes meant that the MVHS data resided in different forms. The NCREL grant provided the occasion and resources for organizing the data in a consistent and accessible database that then enabled the MVHS staff to step back and look at the data in multiple ways previously not possible. We believe that other schools will encounter the need to focus increased attention to data management and data quality, given the increased demands for such sophisticated data handling that have resulted from the No Child Left Behind legislation.
Second, this study has developed simple models for data visualization that are designed to evoke reflection and discussion. Clearly careful thought must be given to what kinds of data and what kinds of data displays are most fruitful for discussions and decisions. Otherwise, the entire data gathering enterprise may expend great energy warehousing data that are only minimally used. As Tufte has emphasized, the ways data are analyzed and displayed determine the informational value of the data.

A starting hypothesis of this study was that a fine-grained look at the achievement data available for students enrolled in MVHS courses would reveal patterns often obscured in tables of data or in summary means based on aggregations at higher levels as is necessary in traditional meta-analyses. This hypothesis was confirmed in multiple ways and in the process of exploring different ways of presenting data, especially using Microsoft Excel, we developed examples that were especially evocative of what we call “data-guided” deliberation. Examples of several of these data displays are included in this report. Through this work, we began to develop the outline of a model for data analyses that addresses the question of What data, what decisions?

Third, we extended our analysis of existing data to look at the automatically generated “course statistics” in Blackboard, the online course management system used for many MVHS courses. These course statistics include counts of student and instructor access to different content areas of each online course by day and time. The course statistics offer a potentially informative, fine-grained look at what is actually going on inside a given course for each individual student. The way Blackboard presents the data, however, does not seem well designed to best inform the instructors or students. We found that by exporting the data into Microsoft Excel we were able to analyze the relationship of measures of student participation in the online environments to measures of students’ final course scores. Although much more detailed analyses are possible, the overall results suggest that in many courses, the simple measure of total student participation is strongly predictive of the students’ final scores in MVHS courses. The importance of this somewhat commonsense finding is that student participation is a measure that instructors could monitor regularly, especially during the first few weeks of a course, in order to intervene promptly when a student appears to be disengaging from participation. The more general result of exploration of Blackboard’s course statistics is a fresh set of recommendations for how such data might be better organized and effectively displayed for students as well as instructors.

Fourth, the analyses highlighted how the wide diversity of types of courses (from credit recovery to advanced placement, for example), subject matter, grade level and purpose all influenced student performance in courses as indicated by attrition and final score. Especially important for understanding the frequent finding of no significant difference is our finding that in some courses the distribution of final scores was bimodal. A cluster of successful students had final scores ranging from 70 to 100 percent and a cluster of students dropped out or received low, failing scores. This distribution suggests that the average score in any given course tends to obscure the more differentiated reality underlying the mean. Data displays that present the viewer with the distribution of final scores within each course serve to invite discussions and decisions on the performance of specific students in specific courses, as contrasted with more diffuse discussions based on means across several courses. The within-course variability among students raised interesting questions that emphasize the complexity of the apparently simple question asking whether student performance in online courses differs from student performance in traditional
courses. The wide variation in reasons why students enroll in virtual high school courses may inevitably lead to wide variations in success. For some self-disciplined, high-achieving students, the online course may be an ideal environment for educational enrichment and acceleration, while other students may enroll in the same course because they were previously unsuccessful in a traditional version of the course, hoping that perhaps this different approach might help them succeed. Our study emphasized the need for careful thought about the different clienteles served by virtual high schools and realistic expectations for the patterns of final scores that might be deemed successful in view of the different clienteles and different types of courses offered.

Fifth, the project resulted in a productive model of collaboration between the “outside” consultant and the MVHS staff. In contrast to more common models where an outside researcher gathers data and analyzes it for publication to the general research community, the present project created a consultative relationship with the staff. In this consultant role, we worked closely with the MVHS staff, beginning with the extensive discussions and iterative process needed to organize the data. This close, sustained collaboration had several benefits. By working with the people who most thoroughly understood the data, the consultant gained a much richer explanatory grasp of how specific events or individual cases influenced the patterns seen in the data. This dynamic interaction between outside consultant and inside staff resulted in enhanced data analytic skills for the staff and the consultant. By working side-by-side with staff, we were able to quickly show preliminary displays of tentative results and ask whether the data were accurate. Once accuracy was established, we asked whether these displays were “informative” or thought provoking, and should be considered for discussion and data-guided decision making.

One enduring consequence of the project is well-organized data on MVHS student performance and staff familiarity with and expertise in ways of analyzing and presenting the data. The analyses revealed areas deserving further attention by the MVHS staff that were previously less salient in the data. Thus, one outcome of this work has been a greater shared sense of the importance of attending to data in the organization. A second outcome of the project resulted from the process by which the researcher worked through the MVHS staff to access the data. In some cases, this led to individuals learning to use new data software, such as MapPoint 2004. In other cases, the analyses called for learning more about the features of available software such as Microsoft Access, Microsoft Excel and Blackboard. In this way, the project could be seen as developing the capacity of MVHS to carry out further analyses after the project ends.

Sixth, the process and results of this project have important implications for both the research and the educational community interested in understanding and improving learning in online courses. More than ever before, schools are under intense pressure to become much more savvy about the uses of data. We conclude this report with an argument that extensive research and development is needed to equip school personnel with the ways of thinking and software tools for data analysis that can enable them to not merely meet legislative demands for data on student performance but more importantly allow them to encourage student improvement through more timely data analysis. Online environments offer new opportunities for use of student data in real-time. The current phase of demands for data seems predominantly aimed at sending test score results up the hierarchy to yield large-scale comparisons of mean scores at the school district and state level. While these global measures attract much attention in newspapers and policy realms, what ultimately matters is what goes on between teacher and student during their time together in
a course or a classroom. Data that inform the moment by moment or week to week decisions of teachers and students needs to be at a fundamentally different level of detail and form than a simple bar chart comparing the performance of one entire school district with the grand mean of another school district.

Seventh, the collaborative, interactive relationship between the researcher and the entire staff of MVHS around what types of data analyses were most informative led to several recommendations for MVHS. For example, the bimodal distribution of final scores, coupled with the potential “early warning signs” available in the course management system has led to MVHS establishing a process for checking on levels of participation in the early weeks of each course. The study has also led to plans for ways of sharing the results with MVHS instructors, with an emphasis on how they can use data available in the course statistics more effectively. The project also led to recognition of the need for additional data to fully address questions raised by these analyses. For example, MVHS provides schools with students’ final scores and the schools have responsibility for deciding what grade to assign to the students’ transcripts. To date, MVHS has not gathered data on what grades are assigned. Such data could serve as a valuable point for discussion with school personnel. Another action following from this project is a procedure for having MVHS staff make more timely use of the course statistics in Blackboard to monitor progress in courses. Future actions relating to course management seem likely to grow out of the increased attention to data resulting from this project.

Finally, we conclude by encouraging the development of a conceptual model for representing data adapted to the specific needs of individuals and organizations through graphs and real-time data dashboards appropriate to the range of decisions available to these players in the educational community. Our project has provided a few initial examples of how data from this specific virtual high school contributed to informed discussion and decisions. Perhaps only when we have developed ways of gathering and displaying data that are meaningful at the teacher and student level will the kinds of high quality data required for more general meta-analyses become available. In the interim, a richly detailed understanding of the dynamics of success in a specific course in a specific context seems like a sufficiently ambitious proximal goal for answering the questions of greatest interest to a student, teacher, parent and administrator, “How good is this virtual course?” and “How might we make it better the next time it is offered?” These questions seem unlikely to be informed by data aggregated at the state level or represented as a single number in a meta-analysis, no matter how carefully it is carried out.

Meta-analyses comparing online with traditional courses have done the educational and policy-making community a great service by giving what seems likely to remain the most honest and accurate conclusion, “There is no significant difference on average between online and traditional courses.” Furthermore, the wide variability within each type of course suggests that every interested party should seek to understand whether any given course is effective and why, irrespective of what might be true of the distributions of all courses. The present study has shown there is much to learn by examining in thoughtful detail data on student performance in specific courses while the courses are in progress, and as soon as possible after a course is complete. The closer in time to the actual teaching and learning that data are explored, the greater the potential benefit to the students and teachers from whom the data are obtained.
This work has suggested several directions for further research. The analyses and insights summarized above have shed considerable light on the sources of variability of student participation and performance in MVHS online courses. These analyses are, however, only a start toward understanding how to apply what we have learned to improve student learning. By virtue of the skills gained and shared sense of the value of data for improving practice developed in this project, MVHS is positioned to pursue several promising avenues for research and development. The vision presented above for how the real-time data gathered in Blackboard’s course statistics could be made much more informative and motivating for students and instructors is one example. The analyses in this report also can serve as the starting point for serious, collaborative work with instructors on how instructor participation via email or discussion boards affects the quality of a course. And the prototype for an administrator’s dashboard invites further development and refinement. MVHS is now well positioned to pursue further research on uses of data to guide decisions and improve online learning. The report concludes with a discussion of some promising topics for future research.
Toward a Deeper Understanding of Student Performance in Virtual High School Courses:  
Using Quantitative Analyses and Data Visualization to Inform Decision Making

Introduction

During the spring of 2005 the Michigan Virtual University (MVU) received a grant to carry out an intensive analysis of available data on performance of students enrolled in the Michigan Virtual High School (MVHS). The purpose of this investigation was to determine whether a fine grained look at this data would help shed light on the reports of no significant difference between online and traditional courses in recent meta-analyses. In addition, the study proposed to develop tools for the visualization of data on student performance.

A Look at Michigan Virtual High School

Online learning has grown rapidly during the past decade at all levels, including K-12. One of the major developments has been the establishment of online “virtual high schools” by many states. The MVHS, as one example, was established with funding by the State of Michigan in 2000 (Public Act 230 of 2000). Since that time MVHS has grown rapidly in terms of number of students served and in the variety of types of courses and “clienteles” for these different courses.

Details on the history and growth of the MVHS are described in a report submitted to the Michigan Department of Education in April 2005 (Michigan Virtual University, 2005), available online at http://www.mivhs.org. Among several services offered by the MVU, the MVHS has provided online courses to over 7,000 students in nearly 400 high schools in Michigan with a total of over 20,000 course enrollments. As such, MVHS is one of the largest virtual high schools in the U.S.

Unlike some virtual high schools, MVHS does not grant course credit or award diplomas independently, but works in partnership with local and intermediate school districts, which award credit or diplomas. MVHS reports the percentage score a student received in an online course to the school and the school decides the amount of credit to award to the student and what grade to assign for the course.

MVHS serves a wide range of students, with the primary clientele being public and private school students in grades 9-12, who typically have an assigned period at school for each course in which they enroll, although they also can work online nights and weekends. Other clientele include home-schooled students, adjudicated youth in institutions, homebound students, students with learning disabilities, and gifted and talented students. This mission of MVHS of serving such a diverse clientele is one of the valuable contributions of the organization, of course, but in the context of this study it is worth emphasizing that this wide variability among students enrolled in MVHS courses makes generalization about student learning in online courses difficult.

Enrollment in MVHS courses has grown rapidly over the past five years, rising from 676 enrollments in the 2000-2001 school year to 5,277 enrollments during 2004-2005. Concurrently,
the number and variety of courses also increased during this period. Similarly, this period has seen significant changes in the funding and tuition arrangements for student enrollment, as well as changes in software vendors for courseware. Details are available in the April 2005 Report to the Michigan Department of Education on the Development and Growth of the Michigan Virtual High School 1999-2005.

Context of the MVHS Study

This study is one of several studies focusing on virtual high schools that were funded by the North Central Regional Education Laboratory (NCREL). These studies were, to varying degrees, stimulated in part by the increased emphasis by the U. S. Institute of Education Studies on research that could be considered scientifically rigorous. In addition, the rapid growth in enrollment in online courses led to increased interest in research on the effectiveness of online learning. Recent meta-analyses comparing learning in online versus traditional courses also contributed to the context for this study of the MVHS.

Meta-Analyses of Online Versus Traditional Courses

In view of the rapid growth of enrollments in online learning, the keen interest of various stakeholders in knowing whether students in virtual high schools are learning as well as they are in traditional classrooms is understandable. Several meta-analyses have sought to tease out an answer to this question.

One of the early and widely cited analyses was done by Russell, who collected 355 studies reporting no significant difference in achievement between online and traditional courses (Russell, 1999), http://www.nosignificantdifference.org/).

Noting several weaknesses of the Russell study, Bernard et al. set out to do a more careful and statistically sophisticated meta-analysis, selecting 232 studies between 1985 and 2002 on student achievement, attitude and retention in online versus traditional courses (Bernard et al., 2004). After exactingly careful analyses, the authors concluded, “Overall results indicated effect sizes of essentially zero on all three measures and wide variability. This suggests that many applications of distance education outperform their classroom counterparts and that many perform poorly” (p. 379). “We found evidence, in an overall sense, that classroom instruction and DE (distance education) are comparable, as have others. However, the wide variability present in all measures precludes any firm declarations of this sort” (p. 416). The authors state, “The most important outcome of the overall effect size analysis relates to the wide variability in outcomes for all three primary measures. While the average effect of DE was near zero, there was a tremendous range of effect sizes (g) in achievement outcomes, from -1.31 to +1.41” (p. 406).
In another meta-analysis Cavanaugh et al. focused on recent studies of K-12 online learning (Cavanaugh et al., 2004). They found that distance education can have approximately the same effect on measures of student academic achievement when compared to traditional instruction. The Cavanaugh et al. report is especially relevant to the current work because it reports on recent studies of K-12 online learning published between 1999 and 2004.

Three findings are especially noteworthy in the Cavanaugh et al. meta-analysis: zero average effect size, wide variability in effect sizes, and a lack of variability of performance due to plausible influences. Specifically, based on 116 effect sizes from 14 Web-delivered distance education programs, they report “distance education can have the same effect on measures of student achievement when compared to traditional instruction.” Or in statistical terms, the mean effect size was essentially zero. As important as the mean effect size, however, is the large variability in these effect sizes, ranging from -1.2 to +0.6. Finally, they analyzed these studies for effects of twelve plausibly influential factors such as content area, grade level, instructor preparation, and so on, but found none of these factors was significantly related to student achievement.

These three findings (no mean difference, wide variability and the inexplicability of that variability) suggest the need for a deeper analysis of what is going on within these courses if future studies, experimental or otherwise, are to reveal more than no significant difference.

**Statistical and Practical Implications of Large Variability**

Several things need to be said about the implications of these repeated findings of such large variability. First, the variability is almost certainly true variability and not mere statistical unreliability, and as such this within-course variance deserves careful attention in research on online learning. From a statistical perspective, when the variability is large, the chance of detecting mean differences is small.

Meta-analytic techniques compare two means, taking into account the variability around each of the means. The typical estimates of variability around the mean are based on presumptions of some degree of a normal distribution. But typically the data available for meta-analyses do not provide detailed data on the actual distribution of the scores from which the standard errors are computed.

From a practical perspective, if the variability is large, then even if an elegant, large sample meta-analysis found a statistically significant difference in favor of one or the other modes of instruction, the implications for practice would remain unclear. Just knowing that the variability is likely to be large should alert teachers and administrators to the need to pay close attention to available data on performance in each course while it is in progress and as soon as it is complete. One need not wait several years for a meta-analysis to tell whether a specific course was in substantial need of improvement.
The Present Study

A primary purpose of the present study is to look in detail at the distributions of achievement scores within courses offered through the MVHS in order to seek a deeper understanding of the variability remarked upon in recent meta-analyses. By systematically delving into these mysteries and by careful analyses of what is known about the variations in performance of individuals in online courses, we seek to shed light on the variability in effect sizes. What has been treated as error variance in meta-analyses — that is, individual variation — is the focus of our analyses. We would note that the variability in achievement of individual students within specific courses is the most important focus of students, teachers, school counselors and parents. Both policy and practice, as well as future design of experimental comparisons, can be better informed by a detailed, quantitative and qualitative analysis of individual performance.

Conceptual Background for the Analyses

Overview of Analyses

The process of analysis reported here involved four phases. The first phase was getting the existing data on student performance into a well-organized form for analysis. The second phase was some descriptive analyses of the enrollments and student performance for major course types over the five years of MVHS’ existence. During the third phase we focused in depth on student performance in online courses offered by MVHS during fall 2004 and spring 2005. Finally, we examined the relationship of student achievement to student activity as automatically recorded in the course statistics for courses offered in Blackboard (http://www.blackboard.com), the course management system used for most MVHS courses.

These analyses proceeded in tandem with explorations of the literature on data visualization, data-driven decision making and data dashboards. Throughout this process we emphasized developing data visualizations that could inform, and illustrative examples of using MVHS data to guide decisions, rather than attempting to carry out a comprehensive analysis of the large number of courses offered by MVHS.

Visual Display of Quantitative Data: Toward Representations that Speak to Stakeholders

One important question concerns what kinds of data analysis have the greatest potential value to stakeholders such as policy makers, teachers, students and parents. Edward Tufte has written extensively about the importance of the visual display of quantitative data (Tufte, 1990, 1997, 2001). His recent analysis showing how the Challenger disaster could have been prevented if the data (on damage to rubber seals) had been presented in a visually interpretable way has received widespread attention (Tufte, 2003).

Academic researchers may tend to underestimate the need to tailor the presentation to different audiences. For example, the typical table in a meta-analysis may communicate well to scholars, but may not adequately engage other viewers. Dickson (1985) describes ways of making software more thought-provoking by juxtaposing representations in different symbol systems. A rich representation of data would also try to move the dialogue beyond the journalistic question:
“What’s the bottom line? Is online learning good?” to more thoughtful explorations of questions as “What kinds of online learning, for what kinds of learners, under what conditions?”

A detailed exploration of individual variation in course performance may be more useful to teachers or parents, for example, than an assertion that an analysis of 200 studies found, on average, no significant difference. This is not to underestimate the value for policy makers and researchers of meta-analyses or to devalue the diligent, exacting work such scholarship entails.

Our purpose here, in contrast, is to approach the task of gathering, analyzing and presenting data with a view to its being useful to multiple stakeholders led to an emphasis on communicating the results of our data analyses in visually informative ways. This approach may be thought of as “reverse engineering” as a way of continually asking the question: “What kinds of information in what form would be of most value to the research and educational community?”

Data-Driven Decision Making in Real-time: Potential for Virtual Courses

While the work on organizing the existing data at MVHS was underway, we began a review of the growing literature on “data-driven decision making.” This review raised important questions and yielded insights regarding the multiple meanings assigned to various terms in the broad area of data-driven decision making.

Although everyone in education is aware of the pressures on schools and teachers resulting from the demands for extensive testing of students and reporting of the results of these tests, what may be less well known among educators and educational researchers is the rapidity with which the corporate sector has moved to offer solutions for school systems to these demands for increased gathering and reporting of data. Of course, at some level such developments are not surprising, but the extent and variety of the for-profit, corporate offerings being developed and marketed to schools may not be well known by everyone interested in uses of data in K-12 schools. Some awareness of these commercial products is fundamentally important for the work of this project, but we believe is important for others exploring uses of data for systematic improvement.

Given our goal of creating new ways of making pedagogically and administratively fruitful uses of data, an exploration of systems developed for the business community can provide examples and food for thought for the educational community. Not only is business better able to make the kinds of capital investments necessary for the development of high quality data analytic tools, but also the competitive environment in which businesses operate places high demands for real-time data presented in ways to enable real-time decision making. The educational world tends to lag in gathering and making data accessible in real-time. A common example in education might be that of publishing data on last year’s reading scores. Any company that tries to operate with year-old data in today’s global economy would likely go out of business. Attention to the potential for gathering real-time data is especially important for online learning because, unlike a regular classroom, students in online classrooms generate real-time data through the process of key stroking and clicking. Real-time data can be used in courseware systems to improve student performance.
Examples of Other Work on Data-Driven Decision Making

Although an extensive review of other work on data-driven decision making (D3) is beyond the scope of this report, a few examples may be of interest in acknowledging the work of others and in placing the present study in the context of this national movement toward increased use of data in education. Although a certain level of interest in data has been around for a long time in education, the federal No Child Left Behind legislation has led to an explosive growth of emphasis on data in education.


Several other concepts share elements in common with data-driven decision making, each with different nuances and discourse communities. Among the widely and differently used terms worth noting are key performance indicators, academic performance index, critical performance indicators and sentinel events.

For example, the California Department of Education has defined an “Academic Performance Index” (API) that yields a numeric score ranging from a low of 200 to a high of 1000, with a target for all schools of 800. Note that this API is a “score” for an entire school. Other systems, such as the System Performance Review and Improvement program in the Oregon Department of Education focus on special education students, with “critical performance indicators” (http://www.ode.state.or.us/search/results/?id=253).

Data Dashboards and the “Balanced Scorecard”

In many ways, the business community is ahead of the educational community in its thinking about uses of data. The concepts of a data dashboard and balanced scorecard seem to have gained some currency in the business community as tools for representing data. The balanced scorecard is most associated with the work of Robert Kaplan and David Norton (Kaplan & Norton, 1996, 2001, 2004) at the Harvard Business School (http://www.balancedscorecard.org) The application

“Imagine entering the cockpit of a modern jet airplane and seeing only a single instrument there. How would you feel about boarding the plane...?”

Robert S. Kaplan & David P. Norton
The Balanced Scorecard: Translating Strategy into Action
of this way of organizing data is being adapted now for use in D3 in the educational
realm, reflecting the increasing influence of thinking of schools through the lens
of business.

Sentinel Events: An Example from the Medical Profession

A concept also worthy of attention from the educational community is that of a sentinel
event, a practice used in the medical profession. The essential criteria defining a sentinel
event include the notion of putting in place a system of alerts triggered by serious events.
These sentinel events lead to immediate action and careful analysis, with lessons learned
being shared quickly with the wider healthcare community.

One theme of the present paper is the importance of enabling educational institutions -- in
this case an online high school -- to develop systems of alerts that enable a rapid response
to improve the prospect of student success in a current course, while collecting such data
over time and from courses, in ways that foster a climate of data-guided decision making.
Early identification of a student who is at risk of not succeeding in a course is as
important to the education profession as a patient having a severe reaction to a new drug
is to the medical profession. Information about the sentinel event policy of the Joint
Commission on Accreditation of Healthcare Organizations is available at
http://www.jcaho.org/accredited+organizations/sentinel+event/sefacts.htm. Perhaps one
element of accreditation review for K-12 schools should be evidence that the school has
put in place alert systems on the educational health of its students, along the lines of this
sentinel event policy.

Creating a Climate for Data-Driven Decision Making in K-12 Education

The educational community has also been actively developing tools for data analysis. For
example, the North Central Regional Educational Laboratory offers several tools for data
analysis (http://www.ncrel.org/datause). Similarly, the Apple Leadership Exchange
provides access to a resource titled, Leadership Practice: NCLB - Using Data for Results,
that was developed by the Southwest Educational Development Laboratory
(http://ali.apple.com/ali_sites/ali/exhibits/1000492). Closer to home, the Michigan
Association of School Boards has emphasized the need to create a culture of
accountability (http://www.masb.org/page.cfm/823). MASB describes “Data to Results”
(http://www.masb.org/page.cfm/422) as a “practical data-driven approach” to increasing
school achievement, stating that this approach “starts with the numbers.” Work on data
design for education could be enriched by drawing on the analogous concepts used in
business and healthcare.

Major Corporate Providers of Data on K-12 Schools

Major vendors are competing for the school data reporting market. For example, Standard
and Poors has been active in many states, including Michigan. At the Website, under the
proprietary name of School Matters (http://www.schoolmatters.com) one can with a few
clicks compare states or schools, identify outperforming schools, and so on. The site
seems to have an inordinate fondness for bar charts comparing two entities and would seem to be most useful in influencing a family’s decision about where to buy a home based on the demographic and academic metrics provided by the Website.

**The Center for Educational Performance and Information**

Michigan’s Center for Educational Performance and Information (CEPI) is charged with gathering and reporting data on Michigan’s K-12 schools ([http://www.michigan.gov/cepi](http://www.michigan.gov/cepi)). Many of CEPI’s reports are primarily in the form of numbers in tables aimed at school building and district comparisons. CEPI, like Standard and Poors, and many others enter the market report data at the aggregate level, not at the individual or classroom level. CEPI is leading the effort in the state of Michigan to bring together data about individual students that currently reside in separate databases by assigning each student a Unique Identification Code within a Single Record Student Database. This is a major effort to create a database that would enable for the first time accurate tracking of student progress and such essential metrics as high school graduation rate. As these databases are being developed, perhaps the examples in the present report might be of some value in shaping discussions of how data in these large databases might be made available in more useful ways at the school level.

In contrast, companies that are beginning to market Web-based instructional packages, such as SAS in School ([http://www.sasinschool.com](http://www.sasinschool.com)) are giving considerable attention to ways of reflecting data in real-time to inform the individual learner as well as provide data on each student and the class in a form that is designed with the teacher in mind. This perspective more closely corresponds to the focus of the present project.

**Home Grown Solutions or Commercial Packages for D3?**

One issue facing school districts, especially small school districts, is whether to buy into one of the commercially developed systems for data management, analysis and data warehousing or whether to use low cost tools like Microsoft Excel to meet at least some of the requirements for data. Of course, these are not mutually exclusive. The present project has focused on developing ways of representing the data available at MVHS to inform the administrators and teachers. We have sought to make full use of Excel’s powerful features, always asking what is the optimal use of data for what purposes.

**Software for Data Visualization: From Access to Excel**

MVHS is in the process of moving to a new system for handling student enrollment and other data that will make future analyses much easier. The work in this project has helped plan for more seamless articulation of this system with a Microsoft Access database and exporting of data from Access to Excel in order to make use of Excel’s data visualization features. Other relatively low cost and easy to use software such as Microsoft MapPoint and xCelsius offer additional tools for exploring and visualizing data.
Considering an imaginary scenario describing how a teacher might make use of the system being put in place in North Carolina at a projected cost of $200 million dollars by 2010. In the recent Education Week’s *Technology Counts*, David Hoff writes,

> “In every Catawba County, N.C. public school, teachers can sit down at a computer and find out just about everything they need to know about their students. They can check grades from current and previous years, scores on state tests, and other student-achievement data. They can track attendance, family history, and other social factors that could affect their students’ test scores. They can read health records, and check up on learning or physical disabilities that may need to be addressed in the classroom. ‘It gives us all the information that we have on a student,’ says Ann Y. Hart, the superintendent of the 17,000-student district, located about 60 miles northwest of Charlotte. ‘The more information you have and the easier it is to access the information, the easier it is to serve the student’” (Hoff, May 5, 2005, p. 12).

This lengthy quote, admittedly a journalist’s imaginary scenario, deserves careful reading, as it conjures up one possible vision of the data drenched world toward which the current emphasis on testing, accountability and data may be taking us. This vision is troubling in several ways. First, a profoundly important issue for society is that of privacy in an era of massive governmental and private databases. Do you as a parent really want the teacher of your child to be able to check up on your child’s family history, health records and grades from all previous years?

Second, are teachers prepared to use such data for the good of a student? Trotter, writing in the same publication, comments on the importance of data coaching for teachers.

> “You can’t just give data to people and expect them to look at it the way you do as a trained analyst,’ says Catherine McCaslin, the supervisor of research, program evaluation, and assessment for… Beaufort County schools in South Carolina” (Trotter, May 5, 2005, p. 17). Even with extensive training, might not the negative impact on children due to stereotyping or finding previous grades lowering a teacher’s aspirations for a child occur more often than positive impact?

Third, the cost-benefit ratio for creating and maintaining such vast databases should not be underestimated. Just because it can be kept, doesn’t mean that it should be kept. And every variable in such databases needs to be entered, checked, debugged, and so on. The amount of time we have spent in this relatively small project at MVHS with 20,000 cases gives some sense of the magnitude of the task of gathering and maintaining so much data envisioned in the quote over year after year. Michigan is spending substantial money in creating a unified database with far fewer variables than implied in the scenario for the Catawaba County teacher.

Fourth, the vision may be unrealistic with regard to the amount of time teachers have for pouring over so much data. A typical high school teacher may see 150 or more students a
day. There is a big difference between data and useful and useable information. The examples in the present paper are illustrations of the importance of careful thought about how to present data, visually or otherwise, for the purpose of guiding decision making. Only if data are automatically presented in visually compelling, prioritized ways are the data likely to have any impact whatsoever.

Fifth, *What data, what decisions?* Serious thought needs to be given to just what kinds of decisions are possible or desirable based on what data? We have begun work on a conceptual model of the range of decisions possible by different stakeholders. The range of decisions open to a classroom teacher, for example, are significant but constrained by time, money, and policy. The same is true for school administrators. Hard intellectual work lies ahead for the educational community at the core of the mellifluous expression, “data-driven decision making.”

These critiques of the imaginary scenario should be seen as cautionary and not as a rejection of the potential value of the increased attention to data in education. Whether the effects of such investments of time and money in substantially greater amounts of data is positive or negative or useless depends entirely on the thoughtfulness with which such data are gathered and used.

**Summary**

In summary, the MVHS data, through the efforts of the MVHS staff, are now nicely organized, enabling us to move easily into the second phase of this study. And through our review of the literature on D3, we now have a clearer, if much more complex, view of how others working in this area are thinking about representing data. This understanding will guide our efforts as we simultaneously look at the powerful ways administrators or teachers in a small school can make use of Excel’s data analysis and visualization tools. MVHS will also consider the larger picture within which the country’s data-driven decision making will take place. Finally, we will keep asking the question: *What data, what decisions?* Much of what has appeared in the literature has been rhetorical assertions saying that surely good will come from having data to inform decisions. Although intuitively appealing, many of the imaginary examples of individual teachers using data that appear in articles lauding data-driven decision making seem rather unrealistic in view of demands on teachers’ time and options in school. With these perspectives in mind, let us now look closely at data on MVHS.
Analysis of Enrollment and Student Performance Data

Visualizing the Big Picture: Placing the Work of MVHS in Geographic Context

In order to properly understand the analyses in this report, we begin by placing the “location” of Michigan Virtual High School in context. When we use the expression “Michigan Virtual High School,” we need to take time to visualize just where this school is located. Perhaps it is worth taking a moment to ask yourself, what images or presumptions do you have in your mind when you read or hear the words “Michigan Virtual High School?” Although “virtual,” MVHS consists of real students, learning in real places, learning from real teachers, and working with real school mentors and counselors.

The map in Figure 1 shows the location of Michigan schools with students enrolled in MVHS courses during fall 2004 and spring 2005. This visual representation was generated using MapPoint 2004 software. Does this visual inform you or change the way you think about the scope of MVHS and its impact on the state? Look closely at the map. Does anything surprise or interest you?

![Map showing distribution of MVHS students across the state and beyond.](image)

*Figure 1. Distribution of MVHS Students Across the State and Beyond*
Consider another question. Do you find yourself looking especially closely at a specific region of the Michigan map above? Where did your eyes come to rest most frequently? Many people zoom in on the region of greatest relevance to them, of course.

A key principle of data visualization is the importance of looking at different scales and a related principle is the importance of providing rich data that engages the mind of the viewer by enabling comparison with an authentic reference group. As an illustration of these principles, look closely at the Grand Rapids area in Figure 1. One notices that this area contains a concentration of students enrolled in MVHS courses, but the resolution makes it difficult to identify specific schools.

Figure 2 (below) provides greater detail for the Grand Rapids area at specific locations where students are enrolled in MVHS courses.

![Grand Rapids Area](image)

**Figure 2. Location of MVHS Students in Grand Rapids Area**

As you look at this zoomed in view, what kinds of insights and questions does this view invite compared with the previous “big picture” of the entire state? If you imagine yourself as a superintendent or principal or parent living in the Grand Rapids area, this map would seem more likely to engage your interest because you would be able to relate the data to schools you know that would serve as a natural “reference group.” Knowing that a school you are familiar with has
significant numbers of students enrolled in virtual high school courses might well increase your awareness of the potential or course availability at your own school. Similarly, this data visualization might lead you to telephone or visit a nearby school to learn more about MVHS.

At some level, these observations may seem obvious, but the display above reinforces our argument that presenting data in a carefully designed visual display promotes “data-guided” discussion, reflection and decision making. We would argue that presenting the data on student enrollments in the visual display above to the various constituencies for MVHS courses is much more immediately grasped than the same data presented as numbers in tabular form.

Although we present only the example of zooming in on Grand Rapids here, similar data displays can be generated for other Michigan regions and placed on the Web as a tool for communicating with, and marketing to, various constituencies across the state.

**Changes Over Time: Placing MVHS in Temporal Context**

Space and time are the two most fundamental dimensions within which human activity takes place. The maps discussed above place MVHS in its spatial context. Another important context for understanding the MVHS and the analyses that follow is the temporal context. As we all seek to understand online learning broadly and MVHS in particular, we should keep in mind how recently MVHS has come into existence. MVHS planning began in fall 1999 and the period of spring and summer 2000 was one of intense start up activity, with funding and the first significant enrollment beginning in fall 2000.
Enrollment Trends over Five Years: The “Big Picture”

Figure 3 shows the growth in total MVHS enrollments from fall 2000 through summer 2005. An enrollment is defined as one student enrolled by his or her school in one course. If the school enrolls one individual student in two courses, that would count as two enrollments.

One principle of data analysis is to understand data in the context of changes over time, where the most common data visualization is a line chart. The chart shows the rapid growth in enrollments in courses offered by MVHS. The chart also shows a slight down turn in enrollments during 2004-2005. Having this trend line in view provides a visually engaging context for discussing the evolution of funding at MVHS during this period. From 2000-2002 MVHS provided the online courses at no cost, using the initial state funding to jumpstart this virtual high school. After these first two years, MVHS has explored several funding models with a view to moving toward a more sustainable model. During the 2002-2003 school year, MVHS continued to offer a highly subsidized subscription program but introduced a tiered pricing model, which continued with some adjustments to the number of “seats” per district during 2003-2004. For the 2004-2005 school year, reductions in state funding for MVHS made it necessary to reduce the subsidy and introduce a new pricing model. (A more detailed discussion is included in the MVHS report to the Michigan Department of Education, April 13, 2005.) The slight decline for 2004-2005 is likely due to the changes in the level of subsidy, the new pricing model and ever-tightening district budgets. MVHS continues to seek an optimum pricing model.
Virtual Summer, Virtual School

On the theme of locating this virtual school in space and time, we invite you to reflect on your image of the traditional high school. What do you consider a typical school year for a traditional high school? As you thought about this, did you think much or at all about the role of summer school in the educational functions of this traditional high school?

[Figure 4. Enrollment in MVHS Summer School versus School Year Courses]

The Michigan Virtual Summer School was launched on April 2, 2004. This program was developed by MVHS to offer the state's high school students a more flexible option for participating in summer school courses (http://www.mivhs.org/content.cfm?ID=135).

Figure 4 shows enrollment in MVHS summer school courses as a percent of courses during the school year for 2004-2005. Although this slice of the MVHS pie is relatively small at this time, it is worth emphasizing the unique potential for virtual high schools to enrich student learning during the summer. Many people have remarked that our current school year evolved at a time when young people were needed to work on the farm during the summer. Today’s working parents do not need their children to work, but would be delighted if educational opportunities were available to their children during the summer. Many students would welcome opportunities to explore elective areas in online environments and have intellectually engaging conversations with like-minded students online during the summer. And students struggling to stay on schedule to graduate would benefit from the opportunity to earn academic credits and avoid losing momentum over the summer. The analyses in the present report are based almost exclusively on courses taken during the regular school year, but we would encourage everyone interested in the future of virtual high schools to keep this special opportunity for virtual summer school in mind.
Districts and Schools Served as Level of Analysis

Our analyses of enrollment trends so far tell us nothing about the success of MVHS in reaching schools and school districts. The past enrollment trends could reflect increasing numbers of students enrolling from a constant number of enthusiastic schools and school districts. Or instead, the trends might reflect a steady enrollment growth from a variety of schools and districts drawing upon the resources of MVHS. Understanding this distinction is quite important in estimating the present and potential markets for MVHS course offerings. For example, if there is a well-defined set of students needing MVHS courses across districts, enrollments should rise as a function of number of districts enrolling students in MVHS. Other plausible models come to mind. The point here is that more detailed visual displays evoke different kinds of questions for data-guided decision making that are masked in overall trends.

![Change In Districts and Schools Served by Year from Fall 2000 through Summer 2005](chart)

**Figure 5. Change in Districts and Schools Served by Year**

Inspection of Figure 5 shows a steady growth in districts and schools served over the first five years of MVHS’ existence, with a slight downturn in schools served during 2004-2005. The fact that during the first two years the number of districts was almost identical to the number of schools might indicate that small districts, with only one high school, may have been prominent among early adopters of MVHS courses. An alternative explanation could be that for districts with multiple high schools, individual schools tend to be the focus of decisions about whether to enroll students in MVHS courses. Understanding which of these two alternative hypotheses is more accurate could help guide decisions about how to better market MVHS courses.
Beyond Space and Time: Toward a Fine-Grained Look at Recent Semester

The preceding analyses have served the purpose of contextualizing the MVHS in space and time, as well as having served to illustrate the value of data visualization as a tool for evoking questions that might serve to guide decision making. We turn our attention now to a detailed look at some of the most recent data on MVHS course offerings and we shift our attention to measures of student performance as contrasted with measures of enrollment. As we shift from analyses of enrollment, which is especially important for financial purposes as well as highly salient for policymakers as they deliberate on budgetary decisions related to MVHS, and to analyses focusing on student performance, the audience for these analyses also shifts. Teachers, students and parents, for example, might have relatively little interest in MVHS enrollment figures; but these stakeholders would likely be quite interested in measures of student performance.

In the following analyses we will focus on MVHS course offerings during spring 2005. We will begin by providing the “big picture” of the types of courses offered during spring 2005. Next we will look at the average student performance by type of course. Finally we will look at different ways of displaying data on student performance in specific courses with a view to informing discussions and guiding decisions about these courses.

At some risk of being repetitive, we would remind the reader that the present study arose in part from the puzzlement of many at the repeated finding of no significant difference between online and traditional courses. A common observation and explanation offered for this finding is that the variability within these comparative studies was quite large. The approach taken in this study has been to drill down into the data in order to carefully examine the variation across types of courses and then to look even more deeply at specific subject matter areas and then within that at specific courses.

As you will see, the MVHS offers a wide variety of courses in terms of content, vendor, grade level and target audience. Any attempt to generalize across such diversity without first understanding the variation across and within these categories would seem unwise. Indeed, if the variability turns out to be large across such categories as subject matter, grade level, and so on, then the question of whether in general online courses are better or worse than traditional courses may prove to be so ill-formed as to be unanswerable. Whatever the eventual answer to this question, we would argue that the approach of “zooming in” and deeply understanding student performance, course by course and subject by subject, at least has the immediate benefit of being able to speak accurately about a course or a handful of courses, which is the level of specificity of greatest interest to schools, teachers and students. With this in mind, let’s continue by taking a look at the kinds of courses MVHS offered during spring 2004 and spring 2005.

“Students across Michigan are hungry for the chance to advance their education in high level math and science courses. We can’t afford to let cost, or distance, or the length of the school year limit their desire for learning.”
Governor Granholm
April 2, 2004
Focusing in on Recent MVHS Courses: Spring 2004 through Spring 2005

Courses by Major Subject Matter Area: The “Big Picture” of MVHS’ Curriculum

In view of the fact that MVHS, like other virtual high schools, has only been in existence for five years, it would not be surprising if most people were uncertain as to what kinds of courses it offered. Ask almost anyone what courses are offered at their local high school and they would be able to name the major courses and give a reasonably accurate description of what courses are typically taken at each grade level. Now ask a typical parent or teacher what courses are offered by the MVHS and the response is likely to be much less informed.

Figure 6 shows the MVHS courses offered during spring 2005, categorized into eight subject matter areas. In terms of data visualization, a pie chart is most valuable for giving an overall impression of the relative proportions of an overall domain represented by the separate categories. This chart makes clear that the “core curriculum” of MVHS’ course offerings are at the “core” of the traditional core curriculum of high school: science, mathematics, world languages, language arts and social studies.

Although only a small part of the pie, the three smaller slices seen in Figure 6 are worthy of note, as one of the potential contributions of a virtual high school is to offer courses that might not have a sufficiently large constituency in any given school. For example, the smallest slice of the MVHS pie, a course in art history, might not otherwise be available to students in many schools, especially when budget cuts are forcing many schools to eliminate such elective courses.
Understanding Sources of Variability in Online Courses

One of the primary purposes of the present study is to shed some light on the sources of variance in student performance in studies comparing online with traditional courses. The pie chart in Figure 6 provides one blinding glimpse of the obvious with regard to this variability. Use of an overarching term such as “online learning” implies a degree of commonality sufficient to warrant generalization about the variety of courses embraced by this term. But how similar are online courses whose subject matter is mathematics or science to those whose subject matter is a world language? Clearly, the subject matters, the pedagogies, and the technology affordances available online have widely differing congruences, not to mention the variability in the sophistication of the software design itself. Depending upon implementation, online courses could offer substantial augmentation of a course on learning French by adding audio and video, or not. Similarly, online courses could integrate simulations and tailored testing into the teaching of mathematics and science, or not, depending upon the quality of the software.

Student Performance in MVHS Courses: “Final Score” as Percent

We begin now to look at measures of student performance in online courses offered by MVHS. Before presenting these data, however, it is important to measure the level of student performance recorded by MVHS and the academic grade assigned by the student’s school. MVHS does not assign a “grade” for a student. Responsibility for grading student performance lies with the student’s school and only the school administration can decide whether to give the student credit for a course and what grade to record on the student’s transcript (http://www.mivhs.org/content.cfm?ID=43). What MVHS does record and report to the school is the student’s final score as a percentage that ranges from 0% to 100%. This “final score” is the measure of student performance that will be used in the analyses that follow.

Average Final Score by Major Subject Matter Category

Consider these questions: Are some subjects more difficult than others? Would you expect the average final score in first year French, for example, to be higher or lower than the average final score in algebra? Or AP calculus? Based on personal experience and observation of others in high school, most people would be likely to agree that some subjects are more difficult, or at least as evidenced by the scores typical students receive in traditional courses.

Before looking at the variability of final scores for MVHS online courses, please pause and consider for a moment your own intuitions or expectations about the “difficulty” of different kinds of courses in traditional schools.

In the next analyses we will focus on spring semester 2005, continuing our theme of narrowing the focus and in this case also concentrating our attention on the most recent data. Note that the subject matter categories differ slightly in the analyses that follow from those in Figure 6, health was separated from “science” because health is not typically seen as a core science. And for simplicity, art history and two small courses were combined as “Other.”
Figure 7. Average Final Score by Subject Matter Area

Figure 7 shows the average final score by major subject matter area for all courses offered by MVHS during spring 2005 semester. The average final scores range from 78% for health to 57% for language arts, with world languages and business toward the high end and mathematics and computer science toward the low end. One’s initial impulse may be to nod knowingly or lift an eyebrow in surprise at the relative difficulty for each of these major subject matter categories.

But further reflection raises interesting questions about expectations, comparative data, and the interpretability of aggregate means such as these. As such, Figure 7 deserves considerable discussion. Consider the variability. Clearly, someone attempting to generalize about whether performance in these online courses for these subject matter areas was better or worse than performance in traditional courses for these subject matter areas would need to know two things: First, one would need to know that the numbers and types of courses and students included in the online and traditional aggregates were similar. For example, the mathematics category for each included similar weighting of remedial math, algebra, geometry and AP calculus. Second, one would need to know not just the aggregate mean but the distribution of scores for both the online and the traditional course offerings.

Interestingly, we have made informal inquiries as to whether traditional schools or state level databases made public the distribution of grades by course. As far as we could determine, such data are not routinely available. Thus, a citizen or an educator wondering about the comparability of performance in online and traditional course is not likely to have any clear sense of what is reasonable to expect.
To take a specific example, what do you know about the distribution of final grades in 9th grade algebra at your local high school? What is the average grade? What percent of students in 9th grade algebra drop out or fail? And what is the average grade for students who are repeating 9th grade algebra having failed it once? The absence of access to this kind of distributional data seems likely to mean that our intuitions and expectations for online courses may be based upon underexamined beliefs about what is generally true or reasonable to expect.

Figure 7 also invites reflection upon what is reasonable to expect as an average final percent in online courses. Again, we invite you to introspect a bit. Many of us have a sense from our own education or as parents of what is a reasonable or typical “grade distribution” in high school. Perhaps the personal experience of most people reading a report like this one is likely to be based on having attended better schools than most and perhaps having done better than most students in those schools. Perhaps one might imagine that a “good” online course would have most students earning scores between 80 and 90%, with some students in the 95% area, some in the low 70% area, and a few failures, with the scores approximating a normal distribution.

Two questions need to be raised about such a “typical” or “normal” expectation. First, when we know that the true high school drop out rate is difficult to specify but is by any metric distressingly high, then there are at least a substantial number of students whose average scores are not likely to be around 80%. Second, statistically speaking, a normal distribution is expected to arise when sampling a reasonably homogeneous population. But as we will see later, the kinds of students who enroll in MVHS online courses may come to these courses for quite different reasons and with quite different prior records as students. Some come for AP credit, while others come to MVHS to attempt to pass a course they had failed at their school, commonly referred to as credit recovery.

Suffice for now to say that Figure 7 sheds light on one source of variability in performance in online (and presumably traditional) courses. Specifically, subject matter areas on average appear to differ considerably in difficulty.

Even this generalization understates the variability by subject matter because the mean for major subject matter areas masks the within-category variability. With subject matter apparently a major source of variance in student performance, let us continue drilling further down, looking in greater detail at specific courses within specific subject matter area, then further down to look individual courses. We will see that even the within-category variability varies by subject matter. There seems to be no end to sources of the variability that confounds attempts at meta-analysis.

**Variation within World Languages by Language**

The subject matter category of “world language” courses offered by MVHS includes three languages: French, Spanish and German. (It is worth noting that in a less global era these might have been called “foreign languages” and that now, during the summer of 2005, MVHS is negotiating plans to offer Mandarin online starting in fall semester.) If you took a “world language” in high school, you may have some opinion as to which of these three languages is the most difficult. On average, in Figure 7 we saw that students in world languages received on average a final score of 77%, but does this average hide large variation among languages?
Figure 8 shows the average final score for French, Spanish and German. The difference in average grade is quite small, as is the sample size for each, but at least for now it does not appear that in world language courses offered by MVHS, there are large average differences in final score. And keep in mind that on average, compared to other subject matter areas, these are among the highest average scores for MVHS courses offered during spring 2005.

What might explain why the scores are both generally high and generally quite similar across these three languages? One possibility is that learning a language is inherently easier than learning mathematics. But other sources of variance, as well as relative lack of variance, are plausible. For the present discussion, consider the fact that all three of these world languages are offered online using curricula developed by one MVHS provider, Power-Glide (http://www.power-glide.com). To the degree that online curricula implement a consistent instructional design, the variability across courses by subject matter (and section) might reasonably be expected to be less than if each language course was developed by a different vendor.

Although it is beyond the scope of the present report, it is worth emphasizing that one of the major sources of variability, both in means and in standard deviations, among students and course in online environments is the quality and characteristics of the instructional design. Virtual high schools (and universities, for that matter) differ considerably in terms of their approaches to purchasing or developing their own online courses. In the case of MVHS, some courses were developed locally, and some were purchased from different vendors. And during the five-year existence of MVHS, there have also been several changes in vendors. These changes pose further challenges to any attempt at generalization about student performance in a given course over time.
Variability Among Science Courses: Two Distinct Sources of Variability

We have seen how average final score in some online courses varies by major content area. And we have seen that the variability among the three world languages is relatively small. Let us look now at the variability within online science courses.

Again, we invite you to ask yourself what your intuitions are about the relative difficulty of high school science courses. Which do you think is more difficult: biology, chemistry or physics? You might ask yourself what your intuition about the relative difficulty of these three sciences is based upon — personal experience in high school or access to data about grade distribution in traditional schools? We invite you to consider another question: Would you expect the final score distribution to be the same or different between the first half and second half of a course, whether taught online or in a traditional classroom?

Figure 9 shows the average final score by subject for five major science courses offered by MVHS during one semester. Examination of Figure 9 reveals considerably more variability among these science courses than we observed in our analysis of the three world languages. The data display patterns in Figure 9 may serve as food for thought, if not for data-driven decision making. For example, the most consistent pattern is that final scores in the second course in a sequence (B versus A) is higher than in the first course in a sequence. Interestingly, the usual language of traditional school does not work here. We cannot refer to these as first semester versus second semester courses, the way these courses would normally be offered in traditional schools. MVHS offers the semester-long A and B courses in the same semester. Therefore, such explanations as “end of the school year” or “senioritis” cannot account for this pattern.

Figure 9. A Comparison of Final Scores by First versus Second Course in Sequence
At least two possible explanations come to mind, each potentially testable. Perhaps only higher achieving students continue on to take the second course (or at least the low achieving students fail and cannot continue), or perhaps the population of students stays the same, but the students get better at learning how to learn in online environments. The data display invites such reflection and discussion, which exemplifies the argument in this paper for the benefits of close examination of data.

Two other patterns may be evident in Figure 9. As is common in many educational settings, average scores tend to be higher in elective courses (anatomy and physiology and environmental science). Second, the “hard” sciences (physics and chemistry) appear to be “hard,” in the sense of having lower average final scores. These observations invite the question of whether such levels of difficulty are inherent to the content, a consequence of instructional design, a result of pedagogy or the way things are expected to be.

In terms of the principles mentioned earlier of designing data displays that are optimal for engagement by specific stakeholders, we invite you to reflect on how different stakeholders might lead to different levels of data aggregation and data display.

Imagine that you are a science teacher meeting in a room with the entire online faculty. The meeting begins with bar graphs showing how U.S. students in general compare to students in other countries on science achievement. Next, Figure 8 is shown on the screen. Perhaps if you take some pride in having high standards for your science course, you would take some satisfaction from seeing that science is more difficult than health or business. But then Figure 9 appears on the screen. What would be your reaction? What influences your reaction?

At this point, presumably your level of interest in the data display is heightened because the data represent your own course within the context of a significant reference group — other science teachers. Your urge to understand or explain or discuss is also a function of which science course you taught.

In terms of designing data displays to evoke meaningful, if sometimes intense, discussions, we would suggest that Figure 9 is potentially fruitful because it packs enough, but not too much, detail regarding two dimensions of comparison that have salience for the teachers of these courses, that being the mean final score by subject and the systematically higher scores in the
second course in the sequence. We would argue that it is precisely at this level of detail that some of the most potentially valuable conversations among teachers and other stakeholders could take place. If the five to ten teachers who taught these courses met in a supportive environment to discuss the data, and probably seek more data, the sharing of ideas, explanations and actions could be productive. And at the risk of belaboring the point, we would suggest that data at this level of resolution and contrast is more potent for educational improvement within these courses than data aggregated at the state or national level.

We have discussed Figure 9 in considerable detail. Keeping in mind that a major focus on the present study is the identification of sources of variability reported in meta-analyses, it is worth mentioning that the pattern with regard to the first versus the second course (A vs B) seems to hold in other content areas.
What Do the Bar Charts Fail to Tell You?

We have argued that Figure 9 has many virtues as a display of data, but now consider what potentially engaging data are missing from Figure 9, or more generally from bar charts. As presented in Figure 9, we know nothing about the variability within each of the means shown as a bar. Figure 9 condenses information about enrollment in each course and the distribution of individual students’ scores down to one mean.

But teachers don’t teach average students, they teach individuals. Let us look back at the world language classes (See Figure 8). Using the somewhat limited charting options in Microsoft Excel, we can sort students in a section of another course from a different semester, in order of final score, and look at the distribution of scores as a bar chart (Figure 10). (Note: We deliberately avoid identifying the specific course or semester because our purpose here is to emphasize the general value of different ways of looking at data and not an evaluation of any specific course. As such, attention to a specific course can be distracting.)

![Distribution of Scores in Percent in One Section of a World Language Course 1A (18 Students)](image)

*Figure 10. Distribution of Individual Scores as Bar Chart*

Inspection of the distribution of student scores in Figure 10 reveals some important information hidden within bar charts. Although some students scored well in this course, others scored quite poorly. Indeed, three students ended the course with zeros. These students appear to have quit working in the course, but technically students did not formally withdraw from or drop the course during the first few days. MVHS distinguishes drops, for which funding is returned, and
withdrawals, which occur after the deadline for a drop. Keep in mind that MVHS does not assign grades but only reports average final scores. Some of these students may have decided after a short period of time that online learning wasn’t working for them; and after consulting with their school mentor or counselor, they may have simply returned to a regular classroom.

From the point of view of MVHS, the examination of the distribution of final scores in this class and others has led to a realization that more data is needed if MVHS is to better understand its success. Among the recommendations at the end of this report is that MVHS regularly gather follow up data on the grades schools assign to students based on online course scores, and also follow up on students who appear to withdraw from courses.

From the point of view of understanding the challenge faced by researchers comparing online versus traditional classrooms (or other researchers attempting to aggregate single studies into meta-analyses), the implication of such wide distribution of scores, including three apparent effective or actual withdrawals, for estimating means and standard errors is profound. We would emphasize that these are logical or conceptual problems and not mere statistical ones.

Consider the increased demands for accountability in terms of such measures as high school graduation rates or adequate yearly progress for every subgroup. Now consider the impact on the mean scores for a school that must include the test score for students who “withdraw”… which might simply mean they put down their pencils and quit marking in the bubbles on the test booklet. Averaging such zeros into the school’s average score may be required by law. Or conversely, if a school is caught excluding low achieving students, the mean test scores reported may be challenged.

In the case of virtual high schools, they may serve several quite different populations, from outstanding students seeking advanced placement courses to borderline drop outs. As you consider Figure 10, we encourage you to keep in mind, “How inadequate is the data typically available on such issues?” As our society moves forward with increased testing and the consequences of such testing, we need to develop a much more nuanced and detailed understanding of the actual distributions of scores, and an awareness of the extreme influence of very low or zero scores upon mean scores.

**Comparing Two Distributions in One Bar Chart**

Visual displays of quantitative data become more interesting when meaningful contrasts are presented. Returning to the theme of the consistent difference between the first and second courses in a sequence (A vs. B), and having made the point that a data display that shows only means in bar charts masks the distribution of individual scores that underlie the means, let us look at a bar chart that adds section B to the section A shown previously in Figure 10. As a data display, we would emphasize that these bars are the actual final scores for the individual students.
What does Figure 11 tell us about A and B for this course that would have been masked in a comparison of two bars as in Figure 9?

Figure 11 shows at least for this course, for this semester, class size was smaller in the second course in the sequence and the average scores are distinctly higher in the second course. Because the student numbers in this case are arbitrary, we cannot differentiate well between the hypotheses entertained earlier that B sections were generally higher than A sections because of selectivity or learning to learn online. A paired comparison for students would, of course, be an even more informative data display.

**Pushing Microsoft Excel to Display Distributions**

In this project we have focused on creating informative data visualizations with the charts available in Microsoft Excel, although much more powerful statistical and graphing software is available (e.g., SPSS or SAS). We have done so for the simple reason that Excel is available in every school and on most home computers of teachers, principals and school board members. This widespread availability precludes the need for schools to spend scarce funds for more sophisticated software. But, more importantly, a person feels a sense of confidence, engagement and ownership when working with data that is significant to them. As a scholar once advised in a research methods course, “Keep your fingers *data dirty*,” meaning stay close to the actual details of your study. (This is quite different from keeping dirty data, but that’s another story.)

Excel has more than enough capability for analyzing and presenting data at the classroom, school and district levels. Given that schools are under pressure to create new cultures of data analysis
and reporting, we believe that developing good examples of data displays in Excel to guide discussions and decision making can play an empowering role for school personnel.

That said, there are limitations in the charts readily available in Excel, especially when it comes to displaying distributions. For example, the bar charts in Figures 10 and 11, while informative, are not ideal representations of the distribution of individual scores. Histograms or plots of distributions such as one might see in a statistical package are not readily available in Excel. (Of course, one can purchase add-ins or program Excel to create histograms, but each level of complexity reduces the probable use and sense of ownership.

*Displaying a Distribution with Excel by Plotting Against a Constant*

We have emphasized the importance of displaying data in ways that engage specific stakeholders. For teachers, the most relevant “chunk size” for data is the class, which typically means somewhere between 10 and 30 students. For samples of this size, one can create an informative display of the distribution of scores by creating a scatterplot of the scores against a constant. And remembering the value of incorporating into a data display an engaging level of complexity and relevant reference group, one can create a rather interesting display of several distributions within a scatterplot by plotting different courses against different constants.

![Figure 12. Displaying Distribution of Final Scores in Three Sections of One Course](image)

*Figure 12. Displaying Distribution of Final Scores in Three Sections of One Course*

Figure 12 shows the distribution of final scores for three different sections of one course. Although the display is not ideal where some dots in the display overlap too closely to be easily counted, this figure exemplifies three important points. First, the distributions in these three sections of the same course differ considerably, as does the class size. Second, for purposes of
engaging the target audience for this chart (instructors and administrators), having more than one class distribution in the chart elevates any discussion to more than a critique of a specific instructor offering a specific course during one semester. Third, the distribution of scores in these three courses is consistent with the pattern discussed earlier, in which a substantial number of students in these online courses do quite well, but a significant number of students do very poorly. Again, we cannot tell why they do so poorly or whether the students themselves think they withdrew from the course. But the display in Figure 12 is offers a good starting point for a discussion around a table involving all concerned — possibly including school counselors who could shed light on individual cases.

**Beyond Final Scores: Using Behavioral Data in Online Courses**

Our analyses to this point in this paper have focused primarily on students’ final percent scores in MVHS courses. We now turn our attention to other data available in Blackboard, the course management system used by MVHS for many courses.

**Students in Online Courses Generate Data in Real-Time**

Students in online courses generate a lot of data compared with students who come to class in a traditional classroom. Every time a student shows up online in Blackboard, data are generated. The data created by students online offers considerable potential for understanding and improving teaching and learning. For example, teachers in a traditional classroom monitor student participation as well as possible, but they do not have time to create a database on the number of times each student responds during class. In contrast, Blackboard generates course statistics automatically — every day for every course.

Potentially, this immediate “coding” of data could be one of the greatest strengths of the newly emerging online environments. But, in keeping with the theme of this project, the question is whether the data are displayed in ways that are designed to inform decisions. Unfortunately, the current data displays provided by Blackboard are quite poor. (A recent press release suggests there may be improvements in the course statistics in the new version of Blackboard.)

**Examples of Poor Use of Data**

Our purpose here is not to carry out a detailed critique of the current ways course statistics are handled in Blackboard. We will, however, describe poor examples, in part because we believe that the current course statistics in Blackboard are so poorly designed as to reduce the usefulness of these statistics tools to instructors.

**Blackboard’s Pie Charts: Mixing Instructor Data with Student Data**

Blackboard records every “click” a student or instructor makes in any area of the course, where there are about 23 categories, such as Content, Discussion, Announcements, Gradebook, etc. The instructor can enter the course statistics area and see what looks like a huge spreadsheet with student names down the left and the categories across the top. The cells contain counts of clicks within each of the areas. Most of the cells are zero, others vary substantially. Course statistics
also provide the same data by day, for each month. Such detail can boggle the mind and raises the question, “What decisions would an instructor make in acting on such detail?”

**Poor Use of Color and Poor Organization of Categories**

Good data displays would highlight the important data in ways that inform the observer. Consider Figure 13, the pie chart automatically generated in the course statistics from one course.

*Figure 13. Pie Chart Showing Distribution of Total Clicks in One Blackboard Course*

The pie chart above is shown in the colors and layout exactly as Blackboard generated them, but the display does not show careful thought or especially good taste in color. Although one cannot infer such from the pie chart, it turns out that Blackboard combines student clicks and instructor clicks in this automatically generated chart. In the course shown in Figure 13, for example, it turns out that all of the clicks in *Student Gradebook* were generated by the instructor. Why would an instructor want his or her clicks included in this graphic? Notice, too, how a number of the “slices” in this pie are so small (essentially zero) — because they reflect obscure or little used features. Other slices, like *Staff Information* or *Observer Tools* are not ones that students would have access to.

In short, the pie chart in Figure 13 is an egregious example of what Tufte describes as misuse of data visualization. Let us suggest here only a few illustrative changes that might highlight student activity in ways that would inform an instructor. First, exclude instructor clicks and
perhaps create a separate pie chart for the instructor. Second, eliminate categories that do not apply to students. Third, sort the categories and start with the most frequent at zero degrees. Fourth, group the slices in some order of educational priority and use color in meaningful ways. For example, put the Content area first, followed by Announcements, followed by the slices having to do with communication, including Email, Discussion Board, etc. These are offered simply as examples.

**Failure to Do the Work of Data Analysis for Instructors**

Blackboard does allow instructors to export the course statistics as a file that can be opened in Microsoft Excel. A dedicated instructor, or one inordinately fond of data visualization, could create meaningful charts along the lines suggested above. And perhaps this would be a good way for instructors to delve into how students in their courses are spending their time. Still, everyone is busy and few instructors have the luxury of such explorations. It would seem that the providers of course management systems should devote more thought to the ways in which data could be reflected back to instructors.

**Predicting Success in Online Courses from Active Participation (“Clicks”)**

An interesting finding in our analyses is that a simple measure of a student’s total number of clicks is strongly correlated with the student’s final score in a course. At some level this may seem obvious. And probably in traditional classrooms there is a strong correlation between a combined measure of attendance and turning in assignments and final grade in a course.

Still, this finding can serve as a clear example of how relatively simple data that is generated automatically in online environments could be turned to good purpose with a modicum of thought and a change in perspective on uses of course statistics.

Although the strength of the correlation between total clicks and final score varies by course, in every case we examined it was strong and positive. The correlation remained strong even when we excluded the students with extremely low or zero (“withdrawn”) scores. A rather typical example is shown in Figure 14. This example uses total clicks in the discussion area of this course, but in most of our analyses we could find little evidence that clicks in specific areas were any more predictive than simply the total number of clicks.
Figure 14. Scatterplot of Relationship between Total Clicks and Final Score

Take time to study the scatterplot in Figure 14. The most obvious pattern is the strong positive relationship between total activity in the discussion board and final score. But again, in the spirit of thinking at a richly detailed level, thinking like a teacher, what else do you notice in Figure 14? Ask yourself about the students...remember that each dot is a real student with a name, who is enrolled in this online course. Which ones fit the prediction best? Who are the exceptions?

For example, consider the student who has both the most participation and the highest final score in the course (100%), who might be thought of as the perfect student. But he or she surely spent a lot of time earning this 100%! Perhaps the wiser student is the one who earned the second highest score in the course but who had a quite average level of participation. At the very least we could say that this student was more efficient or had a higher return on clicks. Perhaps this is the kind of perfect student needed in the new competitive global economy.

As for whether the relationship between performance and clicks is exaggerated by the low achieving students, visualize the trend line if you leave out the students with the lowest scores. What would you estimate would be the correlation with these students excluded? Would excluding these students make the conclusion that total clicks predicts success more or less valid as an assertion?
We could, of course, tell you how excluding these six students would change the correlation of .72. But that would tend to divert attention from our main point here. Displaying data about a relationship between success in an online course and a behavior over which a teacher has some influence, with a sample size of the kind teachers work with every day, is thought provoking. It is more thought provoking than a correlation to some other anonymous teacher’s class in a research journal. It would be a relatively trivial matter of programming to create such a scatterplot routinely during a course, between the cumulating points earned and the cumulating clicks made, and a way to engage the instructor in monitoring participation.

**Teachers Teach Students, Not Dots: Include Student Names in Displays for Instructors**

Teachers think in terms of individual students. Consider how a teacher’s engagement with the data in Figure 14 would be affected if each student’s name appeared in small type beside the dot. We will return to this point soon.

**Use the Computer to Inform Students Directly**

But…and this is really important… why design a feedback system in an online course that requires time and effort by the instructor to make decisions and give feedback to some 20 or 30 individual students? Mind you, as educators we are not opposed to this personal touch, and in fact consider it quite important when it occurs. Teachers are busy, so the more the feedback that can be built into the system in a way that engages the student’s mind and heart without requiring mediation by the instructor, the better.

Remember our criticism that course statistics in Blackboard are visible only by the instructor (and quite badly designed for the instructor)? Imagine the following as an alternative use of the plausible relationship between participation (“showing up”) and success. This is not a bad relationship to be encouraging students to believe in, by the way. During the first week of a course each student is automatically sent an email that asks them to visit a page with a carefully worded discussion of the importance of regular participation, couching this in terms of the importance of self-discipline and punctuality. The page presents a scatterplot like the one in Figure 14 and encourages the student to study it the way along the lines of our discussion above and asks the student to think seriously about their work habits. Of course, the page would also have to caution the student that the instructor was well aware that a student might be tempted to just click a bunch of times and intimate that the instructor knew how to recognize such false participation. The students might be encouraged to discuss this issue with others in the course. Now imagine that about three or four weeks into the class, each student is automatically sent an email that says something like, “You might want to know that compared to others who have taken this course in the past, your current level of participation indicates that you are not likely to be successful. Please email the instructor and explain your low level of participation.”

Enough imagining. The scenario we have sketched out is quite simple to implement technically. What is more difficult is to let go of the hierarchical view that the feedback of such data to students must be mediated through the time and mind of the instructor, as well as the apparent belief that it is sufficient to provide the instructor with page after page of detailed spreadsheet data. We can do better than this in the design of online learning environments.
When Are Students Online in MVHS Courses?

Blackboard automatically generates some potentially engaging charts showing student accesses by time of day and day of week, for each course. These charts appear to combine the access of instructors and students, which reduces their interpretability. Still, these charts do invite reflection and discussion. Earlier in this paper we took time to show where MVHS students were located on a map. Now let’s look at where they are in time. Specifically, when do online students log on? Before looking, however, ask yourself about your preconceptions pertaining to when students in a typical online course are online? During the day? Morning or early evening? After midnight? Weeknights or weekends? And if you were designing a course for virtual high school students, what would you consider to be desirable work habits for your students? The figures below show student access by time of day and day of week, for two different courses.

Access by Hour of Day

Figures 15a and 15b show the number of accesses in two different courses by hour of day. Both charts show midnight beginning at the bottom left (00). A better visual design might be to show the school “day” beginning at 6:00 a.m. (06), perhaps.

![Figure 15a. Access by Hour of Day in One MVHS Course](image)
The first thing that may surprise some readers is that most of the access for these two courses occurs during the school day. For MVHS courses, when schools enroll a student, that student has a regular class period during the day when they are expected to be in a computer classroom. The classroom has a teacher present, but this person is not the teacher for the online course.

When people hear the term “virtual high school” perhaps some assume that most students are not attending a school, or that if they are attending a traditional school, the online course is something attended primarily nights and weekends.

Now compare the distributions of access, by time of day, for these two courses. The charts paint interestingly different pictures of how work time is spent. One of the courses has considerable access in the evening hours and into the early morning. Is this by thoughtful design on the part of the instructor? Is this a good thing? Are these late night accesses common to all students or do they represent a small number of night owls, perhaps due to after school sports and/or jobs. The current analysis is made easily available by Blackboard, but does not enable the instructor to explore these questions in the data displayed. But richer data displays could easily be generated in an online course management system that would encourage students and instructors to think about and discuss good study habits for lifelong learning.
Access by Day of Week

The discussion of access by time of day has served to make the point that engaging data displays can communicate information that might not be widely understood, as well as evoke thought and discussion about important dimensions of learning, study habits and context.

Figures 16a and 16b present data on access in these two courses, by day of week. We will not belabor the productive questions that these visual displays may serve to stimulate, except to say that these two courses, like others we have looked at, differ considerably in terms of the amount of work ... accesses... that students engage in on weekends. As a parent or teacher, would you consider it good or bad if a course engaged a student over the weekend? A recent study reported that many U.S. high school students feel that their high schools do not demand enough of them (Janofsky, 2005). Perhaps future online courses and online assignments in traditional courses could extend the teacher’s influence on how students spend their nights and weekends in productive ways.

Figure 16a: Access by Day of Week: Course A
The examples of variability in student participation by time of day and day of week from these two courses provide only a glimpse of the wide variability on these dimensions that are to be found in MVHS courses. With respect to the theme in this report that a fine grained analysis might shed light on variability in performance, we believe that data on time allocation may offer a valuable tool in better understanding how course design and individual patterns of participation affect student success in online courses. An avenue for further research would be to explore whether students’ patterns of work by time of day and day of week also predict success or failure.

In terms of using visual displays of data to evoke reflection and discussion, the displays of student participation by time would seem to be an excellent starting point for a thoughtful discussion with a group of online teachers regarding pedagogical beliefs about the value of encouraging student homework as well as how the design of courses and assignments affect the level of student work nights and weekends. And as discussed with respect to the correlation of clicks with success, thoughtfully crafted displays could be designed to be shared with the students themselves with a view to cultivating in them awareness and reflection on their work habits compared to those of other students.

Figure 16b: Access by Day of Week: Course B
Designing Displays of Course Statistics to Speak Directly to Students

We would like to offer another fundamental critique of the way Blackboard, like most course management systems, gives access to course statistics. The statistics, including such pie charts as the one seen in Figure 13, are only accessible by the instructor and not by the students. Of course, one would not allow students to see the names of other students associated with the statistics. But careful thought should be given to optimal displays of data to motivate and inform the student. The apparent underlying model is that the instructor would find time on a regular basis to check all of this detailed information, notice in the welter of numbers some deviation or salient figure for a given student, and then email the student or take other actions based on the instructor’s decision. We believe this is unrealistic and certainly inefficient.

What kinds of information might be shared with a student? Currently, students can see the grade they received and the average grade in the class on that assignment — numbers only, no charts.

Consider the possibility of giving each student automated feedback on their level of participation in the course as indicated by the number of times they have logged in to the course and the number of clicks and amount of time they spent in expected areas. The proper display of this information could be quite motivating, and as we shall see, quite predictive of performance.
Data Displays Showing Detail on Individual Students: Personal Learning Trajectories

We have emphasized thinking like teachers and developing ways of visualizing data that engage teachers. In our discussion of Figure 14, we suggested that the scatterplot would become even more engaging if the names of students were visible in the display given to the instructor. (SPSS has an option for printing scatterplots with the person’s name beside the data point. Unfortunately, we have not found a simple trick for doing this with Excel.)

We have also emphasized that online environments provide an opportunity to generate data to inform instructors and students in real-time. Figure 17 is an illustration of how an awareness of the importance of participation or clicks, coupled with the importance of attending to individual students by name, combined with generating in real-time data displays that capture the attention of the instructor might lead to improved learning in online courses.

**Figure 17. Trends for Individuals by Name by Month**

Figure 17 shows the trends in total clicks per month for eight students. These pseudonyms represent real students in a real course. The legend on the right shows each student’s final score in the course. Be careful not to confuse the height of the trend lines on clicks with the student’s final score. Again, we invite you to study Figure 17 and consider what questions and ideas this rich description of these eight students evokes in you. Consider Sally and Joan, who received the two highest scores in the class. They both started out very high on participation but by the last month, they were barely clicking at all. Compare Seth, who ended up with an 83% final score. He was clicking along at a steady rate in February through April but then his number of clicks...
surged in the last month. Perhaps he realized that he needed to burn the midnight keys if he was to be successful. Kris is a less upbeat story, starting out with good levels of participation that decreased with each passing month and ending up with a borderline final score. Of the other four, Larry started participating, peaking in April, and then drifted away. Two really never got started. Perhaps they withdrew, but officially they did not drop the class. One could argue that their data should simply be omitted from any report on success in online courses, but as discussed above, such students also show up and drift away in traditional high schools. Under No Child Left Behind regulations, schools are under increasing pressure to not exclude such students from testing and calculation of annual yearly progress. And as we have said before, we believe that some of the puzzling variability reported in meta-analyses of online learning arises out of just such patterns as shown in Figure 17.

Who knows whether early intervention by the instructor, when an unsuccessful student began to withdraw from participating, might have turned them around? Perhaps the instructor emailed Seth and told him to get to work. Our point here is not that our imaginings about these students are accurate or that miraculous things would happen if data displays such as Figure 17 were provided to instructors. But we believe the serious discussions and reflections that such data displays could promote would, over the long term, contribute to the improvement of teaching and learning in online environments.

**Toward Real-time Data to Inform Decision Making**

Online courses offer new possibilities for understanding and improving teaching and learning. If one wishes to gather detailed minute by minute data on student and instructor behavior in traditional classrooms, one must obtain major funding, make videotape recordings with interpretable sound recording, transcribe or code the data, enter it into a computer, and then begin analyses. In contrast, students in online courses are entering and coding their own data… clicks, entries in discussion boards, chatrooms, and email. The potential for turning such real-time data to the benefit of students has only barely been realized.

**Data Dashboards**

The data analyses and data displays presented above needed to be crafted one by one as we engaged in the entire process of data organizing, data cleaning, exporting to Excel, trying different forms of charts and discussing with MVHS staff the meaningfulness of various displays. In this collaborative process we found two informal criteria useful in judging whether a particular analysis and display was a good one. First, if we showed a data display to MVHS staff and the response was, “Interesting. I did not know that.” Second, and perhaps more important in terms of data-driven decision making, was the response of, “Hmmm…we had better do something about that.” In the conclusion section we discuss briefly some of the recommendations and decisions that have begun to emerge from our work.

Although wisdom and careful thought cannot be automated, decision makers need accurate, real-time data presented in easily understood forms if they are to make informed decisions. In the
analyses and discussion above, we have focused on data displays with an emphasis on the teacher or student as the primary audience.

At the same time we began some initial work on the design of a data dashboard for administrators of a virtual high school. (A data dashboard for administrators of traditional schools would have many elements in common, of course.) The growing popularity of data dashboards in the business world was discussed earlier. More information is available in a presentation to the MVHS Advisory Council (Dickson, 2005); (http://www.msu.edu/user/pdickson/talks/mvhsstudykeypoints.htm).

Using Microsoft Excel, we have created a mock up of what a data dashboard for an administrator might contain. One example of such a dashboard is included herein as Appendix B. This spreadsheet would put essential data on the desktop, such as data on numbers of courses, instructors, students, total enrollment, open seats in courses, and so on. Each of the cells in this dashboard are links to other worksheets in the spreadsheet or Web pages where more information or PowerPoint presentations can be accessed. The dashboard would also provide access to trends and distributions of the kinds presented in this paper.

Using Excel’s conditional formatting, “sentinels” could be set to change the text to red when data on students or teachers signaled potential problems. As elsewhere in this report, we have sought to find ways of doing powerful data displays with Excel. The present version of the dashboard requires that the user click once to update the dashboard with the latest version of the data. With XML (extensible markup language) and other technologies, a fully developed data dashboard would be dynamically updated the way RSS news streams appear on many desktops.

In another mockup, not included here, we have begun working with a commercial software program, xCelsius (http://www.infommersion.com) that creates dynamic dashboard tools that enable the user to interact with the display by dragging one variable to see the effect on another variable. Indicative of the rapidity with which these tools and the dashboard concept are being adopted, the University of Michigan Medical School has just announced that it is adopting xCelsius to create a dashboard for its data. In short, we believe the data dashboard concept has considerable potential for data visualization to inform decision making.

**Supporting Online Teachers’ Learning: Data Design for Pedagogical Reflection**

In our discussion to this point we have discussed ways of using data to inform students directly via better use of course statistics and administrators via a data dashboard. But the heart of education is the teacher. No amount of sophisticated design or technology is sufficient without dedicated and skilled teachers to work with students. The present report focused on data on student final scores and student activity in MVHS courses.

Although not presented in this report, we also took a preliminary look at data on MVHS teachers in terms of how many courses they were teaching in one semester, the sizes of courses taught by teachers and so on. MVHS has carried out a systematic program for developing a cadre of over 300 online teachers. Later in the summer MVHS plans to invite the teachers to come together to discuss some of the data and ideas in this report and use these discussions as a starting point for
finding ways to engage the teachers in dialog about ways of improving online teaching and learning at MVHS. Research on what skills and attitudes contribute to the success of online teachers is needed as much as research on predictors of success of online learners. And, of course, these are patterns of mutual influence that cannot be fully understood independent of the other.

Research on online teachers’ understanding and use of course statistics, as well as how teachers’ beliefs and expectations about the value and use of such course design options as email, discussion boards and online testing could be especially valuable at this point in time.

Conclusions

This final report might better be thought of as an initial report, for though we have learned much, we have barely scratched the surface in our search for an understanding of the sources of variability in student performance in online courses.

Many of our findings are presented in the executive summary accompanying this report and need not be repeated. We have shown how different forms and levels of presenting data about this virtual high school affect the value of the data that supports informed decision making. We would suggest that these ways of thinking about data visualization are applicable to traditional schools and courses. And, of course, many schools already have such data analyses available to parents and school board members in printed versions or on school Websites. We would invite the reader, however, to ask whether such detailed data are readily available at their own or their local school. Can a parent or school board member easily examine the distribution of grades for regularly offered courses such as algebra or biology over the past five years? Has the percentage of students taking advanced science and mathematics courses remained constant or increased over the past five years?

We have built an argument for the value in placing any examination of a course in a broad context and then zooming down to a detailed look at data at the level of a specific course, at a specific time, taught by a specific teacher. Our argument does raise questions about the attempt to reach a single overall answer to the question of whether online or traditional courses are better. But the primary point we would emphasize is that careful attention to data about individual students in specific classes gathered and acted upon while the course is in process may improve somewhat the performance of students in that course; also by taking a careful look at

“...the special task of the social scientist in each generation is to pin down the contemporary facts. Beyond that, he shares with the humanistic scholar and the artist in the effort to gain insight into contemporary relationships, and to align the culture’s view of man with present realities. To know man as he is no mean aspiration.”

Lee J. Cronbach
1975
data while it is fresh and fraught with pedagogical import we are more likely as researchers and educators to better pose the question, “What data, what decisions?” And then perhaps, attempts at generalizing across data from a larger set of such courses may actually become more successful.

Data and accountability are hot topics in education, due in part to the reporting requirements of No Child Left Behind. Many of the potential benefits to learners and schools from all this data seems likely to be unrealized. To the extent that the data are extracted from students, shipped far away for processing and refining, distilled down into mean scores masking rich variability, trucked away on optic fibers to data warehouses in locations remote from the schools and children of origin — to this extent much value is lost.

Cronbach (1975), a scientist who studied education, in his presidential address to the American Psychological Association, cautioned about an undue emphasis on attempting to form broadly applicable generalizations. After spending over a decade studying aptitude-treatment interactions and finding that the research revealed this phenomenon to be far more rapidly changing and context bound than he had imagined, he concluded that our “troubles do not arise because human events are in principle unlawful…. The trouble, as I see it, is that we cannot store up generalizations and constructs for ultimate assembly into a network” (Cronbach, 1975, p. 123; see also Shavelson, Eisner, & Olkin, 2002).

In this paper we have shown how complexity in sources of variability in performance in online courses appears everywhere one looks: subject matter, course within a subject matter, course within sequence, and so on. Because of this complexity, like Cronbach, we conclude that there is not now, or likely ever, any prospect for giving a definitive answer to the question of the relative superiority of online versus traditional courses.

**Designing Graphics to Promote Thought and Discussion**

Information conveyed in graphics can communicate complex relationships quickly. But many uses of graphics are designed to persuade the viewer, to present a tightly packaged conclusion. In contrast, graphics can be designed with sufficient care as to invite reflection, discussion, and then perhaps wiser decisions. Such visual displays of data become thought provoking by juxtaposing representations in different symbol systems (Dickson, 1985).

**Return on Investment: Making Data Useful at Its Point of Origin**

Clearly our society is gathering vastly more data, especially test scores, than ever before, and the number of required tests is scheduled to increase. And competition for human attention is intense. In this situation, how information is presented is crucial if the data are to have any value.

Education is far behind sectors of the economy where competitive pressures have ensured that resources and thought have been given to the rapid gathering of data and presenting it in optimal ways. Consider the stock market, the Weather Channel, personal digital...
assistants, and so on. In the competition for people’s attention, tables full of numbers in an appendix to a school board report — however much money and time has been spent tabulating these numbers — will receive little attention. As argued earlier, if data were made immediately useful and meaningful at the classroom or course level, the quality of data sent up for aggregation would be improved.

Effective Models for Valid and Fruitful Scientific Research on Online Learning

The current priority being given to rigorous scientific experiments with random assignment to treatment condition seems misapplied in attempts at comparing such different entities as online and traditional courses. We are not arguing against scientific experiments per se, but we would raise fundamental questions about validity and feasibility of carrying out a controlled contrast of online and traditional instruction. Random assignment would seem likely to create such distortions in representativeness as to render invalid or ungeneralizable any statistical comparisons obtained. For example, if, as we have observed, many distributions of final scores in online courses are bimodal, with a substantial number of students withdrawn, these extreme scores and the large variance make statistical tests weak. But imagine that in a controlled experiment, only students who were certain to complete the online and traditional course were selected and randomly assigned to treatment. Complete data might be obtained but the samples would be unrepresentative of the samples to which such comparison was intended to generalize.

Design Experiments in Online Courses: Continuous Quality Improvement

In contrast, we believe there are substantial opportunities for models for experimental research that focus on within-course variance by planned comparisons among design features. Similarly, one could consider even random assignment of students to two versions of an online course (which is a variation on the within-course planned variation). More important, by focusing on how variations in course design, including levels of teacher feedback for example, and analyzing the data on student participation and performance with the kinds of visual displays shown above, the researcher-as-designer would be able to test intellectually useful hypotheses.

This kind of focus on continuous improvement is the approach taken by companies working to improve their effectiveness in the competitive, global economy. Concepts such as continuous quality improvement associated with the work of Edwards Deming (http://www.deming.org/) and “Six Sigma” (http://www.isixsigma.com/library/content/c020815a.asp) have brought a scientific, engineering approach to improving performance. These approaches use rigorous, controlled conditions to understand the sources of unwanted variability, and as such have much to offer to the educational research community. And online learning environments, unlike traditional classrooms, offer unique opportunities for applying these design experiment tools to improving learning.

Currently, MVHS does not have reliable data that would enable one to know how students' performance in online course compares to their performance in traditional classrooms. One needs to ask, “Compared to what?” Consider two students in an online algebra course. One receives a final score of 85%, another a final score of 50%. Consider further that one of these students has a grade point average of 1.70 while the other student has a grade point average of 3.56 in their
regular school. What would you be able to conclude about the “success” or “effectiveness” of the online course? Suppose you did not know which GPA was associated with which grade? How would this muddle the question of “success?”

The question a meta-analysis of studies based on random assignment to online versus traditional instruction would attempt to answer is whether on average one mode of delivery is superior. Setting aside the obvious problem of validity and generalizability (in practice different populations of students are found in these two domains), the test of mean differences on average would not shed light on which environment is better for individuals. A more statistically powerful design that could look more closely at individuals would be a cross-over repeated-measures design comparing the same students in an online and traditional course. Short of that level of rigorous control, at the least it would seem that students’ grade point averages in traditional classrooms could be used as covariates in attempting to determine the success of online learning, controlling to some degree for student characteristics.

**Intractable Contradictions at Heart of Meta-Analysis and Other Data Aggregation**

This study began with the premise that one explanation for the repeated finding of no significant difference in meta-analyses was the large variability often remarked upon in these meta-analyses. We believed that a much more detailed look at the sources of this variability might then lead to less ambiguous results in meta-analyses. Not only did we find many sources of variability, we also found that the distribution of scores within many courses was far from the normal distributions assumed in meta-analyses. We have discussed the bimodal distribution we observed in the final scores of some online courses, with a significant number of students appearing to withdraw or only minimally participate. It must be emphasized here that there are certainly analogous extremely low achieving or minimally participating students in many traditional classrooms. Interestingly, we could not find any detailed reports on the distribution of grades and test scores among the published reports on the various national and state-level tests for which means are being widely reported and used for accountability.

We would also point out that the comparisons of performance of school children on the NCLB-mandated tests that appear in the newspapers and on state and federal websites usually represent data reduction far less thoughtful and informative than the data reduction required in meta-analyses. Look carefully at data on 4th grade reading scores in such reports. Usually only bar charts showing means, without even standard deviations, and certainly with no mention of possible bimodal distributions in schools where many children do not speak English or just arrived in the U.S. All data reduction involves a loss of information and detail. The question is whether the resulting measure contains a greater measure of information or misinformation. The examination of bimodal distributions in the present study has highlighted the difficulty in reducing a measure of success to a single percent score.

High school graduation rates, which have been widely discussed in the news and educational policy realms, may serve to illustrate the rather intractable contradictions that arise when attempting to reduce scores on groups to a single number. A layperson might
think this is a simple question, but school districts and states have used widely varying formulas when calculating graduation rate. In Michigan, for example, there are 26 “exit codes” for students leaving a high school and these can be variously coded so as to raise or lower the “graduate rate.” Of course, every organization wishes to look good, and when the stakes are raised as they have been in education for such metrics as graduation rate, then the incentive to manipulate the statistics is increased. The recent agreement by many states to adopt a more uniform procedure for reporting graduation rate is an attempt to constrain some of these manipulations.

Notice, however, the consequences for a school’s average test scores on the increasing number of mandatory tests of keeping low achieving students enrolled. The more “marginal” students a school succeeds in retaining, the poorer the school’s standardized test scores will be. The drive to be sure no child and no group of children are left behind has brought with it intense pressure on schools to be sure that every child’s test score is included. Extensive negotiations continue between the U.S. Department of Education and states over the academic conditions for special education students or recently-arrived immigrants.

Space does not permit a full discussion of this problem. We would simply make the point here that the quest for a single number that can be shown in a tidy bar chart to represent the mean score of a group of children on a test confronts this intractable issue of how completely the score represents all children in the target population. If absolutely all children, including those in special education, the severely handicapped, the non-English speaking, and so on are not included, then the mean scores do not accurately reflect the true score, say the true reading level or the true graduation rate. But extremely low scores, such as we saw in some of the bimodal distributions in the present study, exert an extreme effect on the mean for the course. This problem can be made concrete by imagining a typical class with 20 or so students, whether online or traditional. If most of the students earn scores around 75 or 80 percent, but several students essentially withdraw or drop out and thus contribute scores near zero, the class mean falls dismally.

**A Partial Solution to this Intractable Dilemma**

The emphasis in the present study on displaying richer information rather than information reduced to a single number may offer one way out of this intractable interaction between inclusion and mean score, but not one that will satisfy those who so strongly desire to have a single number for every student. We would suggest that our understanding of the true performance of our schools and classrooms would be more accurate and informative if richer data displays of the kind presented here showed in visually compelling ways the distributions of scores and charts showing dropouts and composition of the student population along with the mean scores that currently seem to be the only datum reported up the data aggregation food chain. Our experiences in the present study in looking in this kind of rich detail and discussing the data displays with staff suggests that this kind of information can evoke thoughtful discussion and reflection that can inform decision making for the improvement of student learning.
We will end by inviting you to think for a moment about the schools, districts, or university academic programs you may be familiar with...as a teacher, parent, administrator, or researcher. Do you have at your fingertips data on grade distributions offered in these educational settings you know? Are data on drop out rates and completion rates for university programs readily available on the Web? Do you know with any precision the grade distributions by instructor, course, or types of students? And perhaps most important in terms of whether such gathering such information would be worth the cost, can you think of data displays that might inform discussions in these contexts around improvement of practice? Demands for data and accountability are here to stay. The challenge will be to work to try to make the effort worth the cost by thoughtful attention to how data might improve practice at its point of origin, one classroom and one course at a time.

Opportunities for Further Research

The analyses and insights summarized above have shed considerable light on the sources of variability of student participation and performance in MVHS online courses. These analyses are, however, only a start toward understanding how to apply what we have learned to improve student learning. MVHS by virtue of the skills and shared sense of the value of data for improving practice developed in this project is positioned to pursue several promising avenues for research and development.

Additional research in the following areas could generate valuable data and statistics to inform policy leaders, MVHS and the K-12 community:

- Research on the use of real-time data to inform and motivate students and instructors, including how different visual displays influence the impact of data on teachers, students, and others.
- Research on what makes online instructors successful, such as the impact of instructor preparation for online teaching, as well as how instructor participation via email or discussion on student success.
- Experimental research using planned variations (such as use of simulations, discussion, chat rooms) within courses to rigorously test hypotheses and enable continuous quality improvement.
- Research on the relationship of the virtual high school to the students’ regular school, including such things as comparative success in the two learning environments, mentor’s decision making when encouraging students to enroll in virtual courses, and so on.
- Research on factors contributing to success of students, including access to technology, interaction with instructors, mentor support, and administrative support in schools.
- Research course content, electronic texts, and other content are used and perceived by students and teachers.
- Research on how online courses could meet the needs of special populations.

Through the work on this project MVHS has created an excellent context within which to develop and test these experimental models for understanding and improving learning.
References

Additional references available at
http://www.msu.edu/user/pdickson/talks/mvhsstudyreferences.htm


Biographical Information

W. Patrick Dickson

W. Patrick Dickson is a professor in the College of Education at Michigan State University. He completed his undergraduate degree in chemistry at Georgia Institute of Technology. After four years in the Marine Corps, he earned a masters degree in science education at Teachers College, Columbia University and returned to Atlanta where he taught physics and chemistry at East Atlanta High School. He earned his PhD from the Stanford University School of Education. After a year as a Visiting Scholar at the University of Tokyo, he joined the faculty of the University of Wisconsin-Madison. He moved to Michigan State University where he is currently the Coordinator of the Ph.D. program in Learning, Technology, and Culture. He is also extensively involved in the Michigan State University College of Education Online Master of Arts in Education program, in which he teaches the capstone Web-based portfolio course. He serves on the MSU Instructional Computing and Technology Committee, which is engaged in on-going discussions of the rapidly increasing use of online learning environments at the university.
MVHS: A Data Dashboard Prototype in Excel

A visual display of data to inform decision making

One of a series of illustrations of making use of Excel's features.

Patrick Dickson  pdickson@msu.edu

Updated:

MVHS NCREL Project on Visual Display of Data to Inform Decision Making

Decision Leader's Dashboard

Comparison: Semester 1 Year Ago  Trend

Courses: number barchart
Students number barchart
Instructors number barchart

Next Semester Oriented Gauges and Sentinels

Courses To Be Offered Next Semester
Current Total Enrollment
Current Total Enrollment vs Projected
Courses Projecting Low Enrollment
Courses Without Instructor

Financial Indices

Revenue by Course and Course Type barchart
Schools by Enrollments large, small piechart
Unused Capacity: Trained Instructors bar chart
Growth Potential: State Priorities, Needs ranked list
Enrollments by Location customers map

PowerPoint Presentations and Reports

Report to Advisory Council: December 11, 2009
Survey results
Dynamic Reports.

On the Web, automatically updated via XML, written to make current data visible to all stakeholders.

Notes:

All Data on Dashboard would be automatically updated and current.
Each category would be an active link to a more detailed sheet.

For example, courses page would provide breakdown by type of course, and within that, a list of actual course.
Flag for Follow Up: Noon

Flags can be set to send reminders.

Five Year Trends Regression

Detailed Look at Courses by Category Over Time

Grade Distributions
Grades and withdrawals
Instructor Histories

Notes:

A visual display of data to inform decision making