THE STANDARD DEVIATION

May 2008

Washington Educational Research Association
University Place, WA

http://www.wera-web.org

State Assessments: Our Work Has Changed

By Robin Munson, Ph.D.

Remember the good old days, when the reading, writing and mathematics tests in grades 4, 7, and 10 were the extent of the assessment program? I do not necessarily remember thinking at the time that life was simple, because I was struggling with the complexities of the new No Child Left Behind rules for federal accountability. ("Does the 'A' in AYP mean 'annual' or 'adequate'?") In retrospect, however, those were uncomplicated times. That was when an annual assessment workshop and a few OPSI memoranda could help me understand the state's assessment system sufficiently to explain it to the curriculum director, the superintendent, and even (annually) to the school board. That was when I could complete the SPSS file for that spring's WASL and only revisit it when someone needed further disaggregation. Our work has changed, and it is no longer as straightforward as it used to be.

The imminent graduation of the first class of students required to demonstrate that they have earned a meaningful high school diploma seems a good time to reflect on the ways in which our work has changed. The evolution of Washington's assessment system over the past few years has resulted in changes we, assessment experts at the local and state levels, must enthusiastically embrace.

The stakes have changed. Our work has changed significantly. It has shifted from "simply" providing state, district and school accountability, to adding individual student accountability as well. Because diploma determinations depend on assessment results, our work has shifted from program and system evaluation to accountability for every single student – accountability for mastery of content, communicating the rules and options, and providing the opportunity to learn.

INSIDE THIS ISSUE

The questions have changed. The primary question we assessment directors used to be asked was something like, "How many 10th graders passed the reading WASL last spring?" A sophisticated follow-up question might have been, "What percent of 10th graders didn't test this time?" Now the predominant question is, "How many 12th graders have met the reading and writing standard (through the WASL/WAAS/CAA Option/waiver) during their high school career?" Additional questions include: "Who needs to be in summer school?"; "How many students have never accessed the test?"; "How many students improved their scores through a retake, and by how much?"; "How many students met standard using an alternative assessment, and/or a CAA Option?"

The audiences have changed. The increased focus on individual students means the community of concerned stakeholders has also changed. Every principal, counselor, teacher, curriculum specialist, and communications manager now joins the superintendent and school board in looking to the assessment department for information, clarification, and results. Additionally, individual students and parents have a much higher interest in the tests, the scores, and how the whole system works.

(Continued on next page)…

The timelines have changed. The timeliness of reporting is critical when decisions about graduation, summer school, course enrollment needs, and assessment options are on the line. To expedite reporting to students and their families, timelines for processing the tests have been squeezed at every step. There are no more margins for error in the processing of test results. Timelines that used to be able to tolerate minor glitches and still deliver results on schedule are now extremely unforgiving.

The complexity has changed. The advent of a standards-based graduation requirement has created a new sophisticated network of responsibilities. We've created primary assessments, alternative assessments, options for earning a certificate of academic or individual achievement, and special circumstance appeals and waivers processes. Courses have been developed in several subjects to support the mastery of the standards. Students who don't meet the mathematics standards as measured by our state assessment are required to earn additional credits in math until they have met those standards.

Adding to the complexity, students have opportunities to take the high school assessment multiple times, beginning in 9th grade, and in multiple environments (i.e., on-line programs, alternative programs, high school completion programs, etc.).

Annual workshops are not sufficient, so we've created a weekly newsletter, *Washington Assessment Weekly*, for district assessment coordinators, which regularly contains multiple pages of updates.

The data management has expanded. Longitudinal tracking is paramount when students have multiple opportunities and multiple methods to meet the graduation requirements. Data management practices have shifted from point-in-time single administrations to cumulative longitudinal records. Data reconciliation across test administrations, and across assessment options, is vital though not easy. For example, scores that previously were suppressed for AYP calculations now must be visible to determine graduation status.

Measurement issues are becoming more prominent. The assessment system's primary measurement issues used to focus on the design, scaling, and equating of each year's tests. Now, in addition to those concerns, the measurement system is being asked to:

- address questions of how to measure the progress of individual students across time,
- assess newly developed content standards in a way that is distinct from, yet clearly connected to, current measures, and
- to provide technically sound measurement tools for the diagnosis and monitoring of individual students.

Additionally, the technical quality of the state's tests, and of the approved alternatives, is of more and more interest to in-state stakeholders and to federal officials.

Changes are going to continue. With new mathematics and science standards on the way, new assessment contractors soon to be announced, and subject matter tests (a.k.a. end-of-course tests) on the horizon, more change is ahead. But when I reflect on the change we have managed in Washington in the last few years I am reassured we will continue to successfully address these changes, and am hopeful that, with an eye to quality and fairness, the new changes will result in an improved system. I am energized by the challenges ahead, knowing that I am surrounded by great assessment professionals, both at OSPI and throughout the state, and that our work is enhancing the education of every single student in the state.

-Robin Munson is Director of Student Information at OSPI and meets regularly with the WERA Test Director group. She was for many years with Tacoma Public Schools, most recently as Executive Director of Research and Evaluation.

President's Column

The WERA mission is to improve the professional practice of educators engaged in instruction, assessment, evaluation, and research of the following:

assessing student performance; evaluating programs; conducting and applying educational research; and using data to inform instructional decisions. One of the important ways we carry out this mission is through academic discourse. WERA members can find ample opportunities to discuss educational practice through *The Standard Deviation* articles, by reading educational research projects and white papers, and attending WERA-sponsored conferences.

The Standard Deviation online publication has expanded in the past few years to include thought-provoking summaries of recent research at the state and school district levels. Other topic areas include program evaluation models, tips for data analysis, and book reviews. All of these articles are designed to pique your interest in starting a new conversation with your colleagues about educational practice in our schools and districts.

Another way that WERA promotes academic discourse is through sponsorship of new research projects. This year's grants focus on math preparation for college, sustaining high school conversions, and consequences for teaching English Language Learners. Each research team will present its findings at a future WERA conference. The WERA board has also sponsored white papers as a result of collegial conversations about issues faced by school districts.

These reports can be found on the web page: http://www.wera-web.org/pages/publications.php

One of the best venues for educational conversation and discussion is at WERA conferences and workshops. The annual December assessment conference in partnership with OSPI features both keynote speakers and presenters who challenge our thinking and inspire new ideas in assessment. The spring WERA conference features research and best practice in curriculum and instruction as well as program evaluation and research. These conferences and one-day workshops afford us with the time to examine and improve our professional practice. Thank you for participating in WERA and sharing your thoughts with all of us.

-Nancy Arnold, Ed.D. is Director of Special Programs for Puyallup School District and WERA President. She was a special education assessment specialist with OSPI prior to accepting the Puyallup position.

The mission of the Washington Education Association is to improve the professional practice of educators engaged in instruction, assessment, evaluation, and research.

WERA Services

- *WERA provides professional development through conferences, publications, and seminars.*
- *WERA provides forums to explore thoughtful approaches and a variety of views and issues in education.*
- *WERA provides consultation and advice to influence educational policy regarding instruction, assessment, evaluation, and research.*

Plan to Attend the Winter Conference December 3rd through 5th, 2008

Plans are moving ahead for the 24th Annual Washington State Assessment Conference on December 4-5, 2008, at the Seattle Airport Hilton Hotel Conference Center. This conference is sponsored by the Washington Educational Research Association and the Office of Superintendent of Public Instruction to provide in-depth assessment and research sessions for educational professionals. Please save the dates for both the conference and the pre-conference half-day sessions on December 3rd.

Even more than disseminating information, the WERA board intend for this conference to provide a place for conversations about how assessment dynamically informs instruction and program decisions. Two keynoters have abundant experience and expertise in assessment and teaching:

- Dr. Henry Levin is William H Kilpatrick Professor of Economics and Education and director of the National Center for the Study of Privatization in Education at Teachers College, Columbia University. He has written or edited several books, including *The Price We Pay: Economic and Social Consequences of Inadequate Education.*

- Dr. Terry Bergeson, Superintendent of Public Instruction, will share recent assessment and program updates at the conference, including the implications of revised math standards and the review of science standards for state assessments.

In addition to the keynote speakers, more than 50 breakout sessions are being planned for Thursday and Friday. These sessions are presented by local school district educators, researchers, program directors and OSPI staff. There will be presentations on a number of educational topics designed for teachers, principals, central office staff and high education faculty and students. On December 3, pre-conference workshops offer extended learning experiences in areas of high interest, such as No Child Left Behind, revised math standards, state assessments, and high school strategies for success and graduation. Free clock hours will be available for all days of the conference.

WERA is accepting proposals for presentations through June 15, 2008. If you are interested, you can download the presenter's form at www.wera-web.org/links/WC08_Presenters.doc. Submittal information by email or fax is included on the presenter's form. Conference registration forms will be mailed in September and online registration can be found at www.wera-web.org next fall.

Henry M. Levin is the William Heard Kilpatrick Professor of Economics and Education at Teachers College, Columbia University.

Washington Educational Research Association
WERA/OSPI State Assessment Conference, December 3-5, 2008
Seattle Airport Hilton Hotel Conference Center

Presenter's Proposal Form

Type of Proposal (check one): ___Half-day Wed. Pre-conference ___75 minute breakout session Th-Fr

Name_____ **Phone** _____ **E-Mail**_____

Address_____**FAX**_____

Job Title_____**Employer**_____

If others will assist in this presentation list their name, employer, and title for inclusion in the program:

Complete these lines *only if the presenter listed above is not the person to contact* about this presentation:
Contact Person:_____ Phone _____ E-mail_____
Complete mailing address: _____

Presentation Title:

Brief Description of Presentation: (Please limit to about 50 words, edit carefully, and write legibly.)
☆☆☆*This is for inclusion into the program. An interesting description will attract people to come.*☆☆☆

Primary Audience(s): (Mark with an X for all that apply. This helps us provide balance when preparing the conference schedule.

___Teachers ___Building Administrators ___Central Office Staff ___University Staff
___Research/Evaluation/Assessment Specialists ___Elementary ___Middle School ___High School

Schedule conflict: (Mark with an X only if you are unavailable to present on the indicated day/time.)
___ Unable to present on Wednesday, December 3 in ___ morning; in ___ afternoon
___ Unable to present on Thursday, December 4 in ___ morning; in ___ afternoon
___ Unable to present on Friday, December 5 in ___ morning; in ___afternoon

***Email** this **completed** form as an attachment ASAP, but no later than June 15, 2008 to:*
Nancy Arnold: arnoldnl@Puyallup.k12.wa.us or FAX 253-841-8655.
Call Nancy if you have questions: 253-841-8700.

WERA Annual Membership Meeting

President Lorna Spear called the meeting to order at 12:23 p.m. on March 27. Approximately 150 people were in attendance.

President Spear reviewed WERA accomplishments for 2007-08:

- Efforts to advance a social justice agenda have been made through articles in the "Standard Deviation" newsletter. There was a conference strand included at the December Assessment Conference, and sessions were included within this strand on the program.
- A conscious effort has been made to increase the involvement of higher education staff in WERA programs. The number of conference sessions with a research focus presented from university staff members increased from 9 in 2006 to 24 in 2007. Additionally, a math in-service training by Ruth Balf, professor at the University of Washington, was sponsored by WERA earlier in the year.
- Awareness of current research has been promoted through increasing the number of articles in the "Standard Deviation." The sessions in conference programs that focus on research tend to be attended by fewer people than other sessions. A continuing effort will be needed to encourage more people to attend these sessions.
- Emphasis in the awards and grants area can be seen in the award given by the Assessment Directors Network, and an award for the outstanding dissertation (see detail below). WERA was able to increase the funding of grants from $6720 this year to $16,000 for next year.
- WERA membership now exceeds 550 for the year, which is slightly higher than it was last year.

President-elect Nancy Arnold was asked to announce the WERA award winners. Nancy announced the following WERA awards for the year:

- Outstanding Dissertation Award – to Hilary Loeb for her University of Washington dissertation entitled, "National Board Certification as a Support for Work with Historically Underserved Students: A Case Study of Washington State Teachers" (Hilary was unable to attend conference due to attendance at national AERA meeting.)

- Gordon B. Ensign Award – to Joe Willhoft, OSPI, honoring his long career of outstanding contributions to education in Washington. (Joe was unable to attend conference due to attendance at national AERA meeting.)

- Art Maser Service Award – to Michael Power, Tacoma School District, for his outstanding service to WERA.

- Emeritus Membership – to Sue Shannon, for many years of active WERA participation in planning activities and presenting at conferences, retiring from OSPI.

President Spear returned to the podium, and personally presented the awards to Michael Power and Sue Shannon, who were both present at the conference. She also acknowledged Robin Munson, who had received a WERA award earlier in the year from the Assessment Directors Network.

President Spear reported that WERA's fiscal position is sound, and that a balance has been successfully maintained to self-insure for unexpected future fiscal problems, and pay for ongoing commitments. WERA relies on conference attendance as the primary means of supporting its activities. Conference attendees were informed that printed copies of fiscal and membership data and were available at the WERA registration desk.

President Spear called outgoing Board member Pete Bylsma to come forward to receive thanks and a plaque honoring his most recent three years of service on the Board. Pete was also co-chair of the current conference.

(Continued on next page)...

President Spear announced the results of the recent WERA elections as follows:

- Emilie Hard from the Tahoma School District was elected as president-elect.
- Phil Dommes from the North Thurston School District was re-elected to a three-year at-large Board Position.
- Paul Stern from the WSU SESRC in Olympia was elected to a three-year at-large Board position.

President Spear called Nancy Arnold to come forward, and presented Nancy with a gavel to use in her upcoming year as the new WERA president, which begins on May 1. Nancy thanked Lorna for all of her good work, and presented her with a plaque honoring her service as president.

The meeting was adjourned at 12:38 p.m.

Future Calendar

WERA Items

- 2008 State Assessment Conference,
  December 3-5, 2008
  Hilton Seattle Airport Hotel
- 2009 Spring Assessment Conference,
  March 25-27, 2009
  Hilton Seattle Airport Hotel
- 2009 State Assessment Conference,
  December 9-11, 2009
  Hilton Seattle Airport Hotel

Contact: http://Wera-web.org

Other Calendar Items (Non-WERA)

- School Improvement Through Essential Partnerships: A Framework for Leading Teacher Growth (WSASCD)
  August 18th, 2008
  Todd Beamer High School, Federal Way
- American Evaluation Association Annual Conference, Denver, CO,
  November 5-8, 2008
- WSASCD Annual Conference,
  Spokane, WA,
  November 6-8, 2008
- Creating Active Readers: Applying Research-Based Comprehension Strategies in Grades 2-8 (WSASCD),
  Betz Elementary School, Cheney, WA
  January 10, 2009
- American Educational Research Association, National Council on Measurement in Education, National Association of Test Directors, Directors of Research and Evaluation Annual Meetings and Conferences, San Diego, CA,
  April 13-17, 2009

OSPI Conferences Contact:
http://www.k12.wa.us/Conferences/default.aspx

Websites of Interest to Measurement Folks

http://www.elsevier.com/wps/find/editorshome.editors/Introduction Elsevier's hosting of an ethics site for professional journal editors

http://www.gamequarium.com/measurement.html Measurement games

http://www.onlineconversion.com Convert just about anything to anything else

ftp://ftp.sas.com/pub/neural/measurement.html A 1997 primer on measurement theory

http://www.census.gov/hhes/www/povmeas/relatedsites.html Census Bureau sites for measuring poverty including the West Coast Poverty Center at the University of Washington

Book Review: Transformative Assessment by W. James Popham

Reviewed by Phil Dommes, Ph. D.

In slightly less than 150 pages, Popham provides an informative, succinct, and very readable discussion of formative assessment. He begins, appropriately, with a definition: "Formative assessment is a planned process in which teachers or students use assessment-based evidence to adjust what they're currently doing." Popham distinguishes formative assessment from benchmark or interim assessment which, although potentially useful in marking progress on mastery of important outcomes, have not been supported by research as educationally beneficial. What has been supported, and what Popham promotes, is a classroom-based formative assessment process, leading to teacher decisions to adjust instructional practice or student decisions to adjust learning tactics.

Terms defined, Popham turns to implementation. His first consideration is how teachers know what to assess. Popham's discussion centers on the idea of "learning progressions," or ordered sets of subskills or bodies of knowledge that students need to attain a broader curricular goal. Better known to some of us as "task analysis," the concept of learning progression is a key step in planning assessments.

In subsequent chapters, Popham delineates four levels in the implementation of formative assessment: 1) teacher's instructional adjustments; 2) students' learning tactic adjustments; (3) classroom climate shift; (4) schoolwide implementation. For the first two levels, he systematically lays out a structure for using various kinds of formative assessment for teacher and student purposes. He considers when to assess, how to assess, how to set triggers for the adjustment of instruction, and how to make the adjustments. In the classroom climate level, Popham includes a helpful discussion of the kind of teacher activities that promote a classroom where a culture of formative assessment can flourish. The schoolwide implementation level emphasizes the importance of professional development and, most importantly, the use of teacher learning communities.

In summary, Popham brings considerable clarity and insight to the topic of formative assessment. As indicated in his title, he shows how an informed implementation of formative assessment can truly be transformative for students, teachers and schools. What Popham does not deny, but what may get hidden in his simple prose and stepped-out explanations, is the amount of work and persistence it takes to sustain the approach to formative assessment he describes.

Publication Data:
Transformative Assessment, by W. James Popham, 2008, Association for Supervision and Curriculum Development, Alexandria, VA. Paperback, 148 pages, 22.95 (US) ISBN 978-1-4166-0667-3

-Phil Dommes, Ph.D., Director of Assessment and Gifted Programs, North Thurston Public Schools

More Websites of Interest to Measurement Folks

http://www.unc.edu/~rowlett/units/ A dictionary of units of measurement including **Pfiff**
a traditional unit of liquid volume for beer in Austria. A Pfiff (the German word means "whistle") is a small quantity of beer. Traditionally it was equal to 1/2 Seidel, which would be about 177 milliliters (5.99 U.S. fluid ounces), but in current use it is generally 200 milliliters (6.76 U.S. fluid ounces), which is 2/3 of the metric Seidel in Austria.

http://funnel.sfsu.edu/creep/CreepMap.html
Track fault creep along California earthquake zones.

http://www.fairfaxcounty.gov/dmb/pm_news.htm
Fairfax County, VA newsletter for performance benchmarking, an interesting model to study as District Strategic Plans and School Improvement Plans are reviewed.

The Standard Deviation to be Indexed

By Andrea Meld, Ph. D.

Coming soon to the WERA Website http://www.wera-web.org/pages/homepage.php?page=homepage is an index for *The Standard Deviation*, the organization's newsletter, by topic, type of article, and author. The on-line index is designed to make articles of interest from past issues of *The Standard Deviation* available at the reader's fingertips. The first type of article to be indexed will be the 'how-to' articles for software applications such as Excel, SPSS, Access and others, for the technically-savvy and those wishing to become more savvy. These articles have been written by WERA members, who are district assessment coordinators, assessment analysts and others interested in sharing their tricks and tips. The entire index, which will continue to grow with each edition of *The Standard Deviation*, will be posted at the beginning of summer, according to plan.

We encourage you contribute to *The Standard Deviation*, to "Share what you know, learn what you don't" -Raynald Levesque, SPSS guru.

-Andrea Meld, Ph. D., Assessment Data Analyst, OSPI. She is currently serving as Member at Large on the WERA Board and is on the December 2008 Assessment Conference Planning Committee.

Cool kids from Fawcett Elementary in Tacoma played the marimbas at the Spring WERA Conference.

The Golden Age of Program Evaluation

By Peter Hendrickson, Ph.D.

Directors of Research and Evaluation (DRE) from around North America met in New York City before the annual American Educational Research Association in March. A common thread in the reports from members present was the pronounced tilt to assessment at the expense of program evaluation. Only the largest districts (Chicago, Wake County [NC], St. Paul…) appear to support evaluation services and many good sized districts spoke of lost evaluation capacity.

Keynote speaker Michael Kean, retiring CTB-McGraw policy vice president, spoke of the 1970's as "the golden age of program evaluation" with the explosive growth of federal programs. The education labs took on a focus of program evaluation. Philadelphia had 175 FTE evaluation staff in 1977. A Web search showed only 5 evaluations stamped in SY 2006-07. Graduate student enrollment increased dramatically in program evaluation at the same time.

School district research and evaluation office produced many landmark evaluation studies in those years, Kean noted. The Directors of Research and Evaluation started in 1978 while Pres. Ronald Reagan was attempting to disband the U.S. Department of Education. The passage of Goals 2000 under Pres. Bill Clinton ushered in calls for a national test and NCLB brought in a focus on testing, rather than evaluation. The former Baltimore City Schools program evaluator mused that the backlash against testing is truly against how the data are used. He predicted reauthorization, likely in 2009, which could shift the focus of use of the data to improve instruction.

Administrators who have worked in districts with declining enrollments (and budgets) often point to the need for robust program evaluation to make difficult personnel and program continuation decisions. Expertise in developing, administering, scoring and analyzing results of standardized tests has grown remarkably in the past five years. At the same time, program evaluation skills have atrophied. There is a robust culture, with staffing, in districts of assessment (to the dismay of many) but no parallel culture of program evaluation, only teacher evaluation.

After 30 years, DRE members explored both a redefinition of their identity and named unique purposes for the next decade. Collaborative research or program evaluation of a nationally used program such as Read 180 was discussed as a defining activity. Considerable discussion centered on becoming a Special Interest Group (SIG) of AERA.

-Peter Hendrickson is an Assessment, Research and Evaluation Specialist for Everett Public Schools and served three years as an officer of the Oregon Program Evaluators Network. He edits The Standard Deviation.

Consultants, Contractors -Speak Up

We are preparing an article for the next issue about working with contractors. This will include a list of WERA members who provide consulting/contracting services to schools, school districts, and ESD's in Washington. If you would like to be identified in the article, please send an e-mail to Paul Stern, sternpo@wsu.edu with the following information:

Company Name
Names of Individuals who are WERA members in your organization
Primary e-mail address
Primary Phone number
Description of services offered

-Paul Stern, Ph. D., is Senior Research Associate with the Social and Economic Sciences Research Center at Washington State University/Vancouver. He was recently elected to a WERA Board Member-at-large position.

Test Directors Network Update –A WERA Affiliate

The WERA Test Directors Network will meet Monday, August 4th, 2008 at the Seattle Hilton Airport Hotel to review WASL test scores with OSPI staff. The meeting will be from 8:00 A.M. to 2:00 P.M. Attendees have been asked to be prepared to share local results. Inquiries to convener Bob Silverman at Puyallup Schools. Contact him at SilverRJ@Puyallup.k12.wa.us

Local Statisticians Meet the WASL

OSPI Assistant Superintendent Joe Willhoft and Test Director Yoonsun Lee addressed the Puget Sound Chapter of the American Statistical Association (ASA) to roll out, in considerable detail, the test development process for the Washington Assessment of Student Learning (WASL).

The April 17th dinner meeting in Seattle drew several questions about item selection including classical test theory and item response theory. Unlike many presentations to statewide audiences, the statisticians showed keen interest in the symbolic representations of test theory. ASA members are drawn from the diverse communities of academe, Boeing, program evaluation, medicine and other scientific areas.

Willhoft also addressed recent legislative action impacting future testing including high school end of course mathematics assessments.

Joe Willhoft chats with a ASA member regarding WASL development.

Yoonsun Lee and Puget Sound Chapter ASA Outgoing President Bruce Peterson.

WASL Growth: Exploring Issues in Growth Data

By Don Schmitz

Math and Reading WASL scores for the first time are available for students in successive years. WASL growth scores can be used to measure growth for buildings, grade levels, programs, classrooms, high achievers, and low achievers.

There are four models of accountability (see Figure 1) that Pete Goldschmidt (2005) describes in his "Policymakers' Guide to Growth Models for School Accountability: How do Accountability Models Differ?" This article will investigate a Growth Model using WASL scores from successive years. It will attempt to answer the question, "How much on average, did the same students' performance change?" It will build upon an easily understood model of WASL growth and discuss issues of comparison.

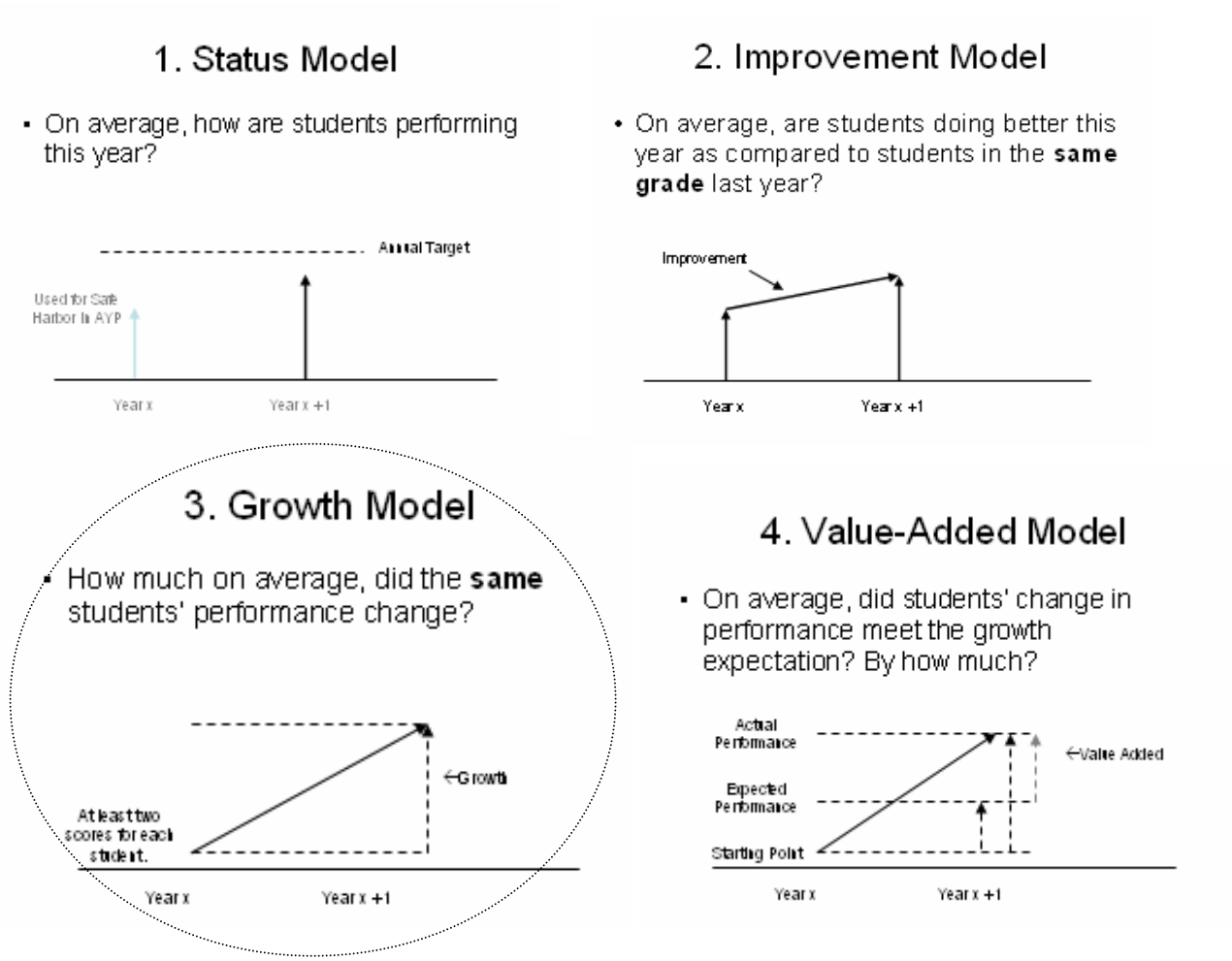


Figure 1. Four Models of Accountability.

WASL results are available for grades 4 through 8 for 2006 and 2007. For the first time we have students taking the WASL in successive years in successive grade levels (see Table 1). Unfortunately, the WASL is a different test on a different scale for each grade level and each subject so comparisons across grades are inappropriate.

(Continued on next page)...

A vertical scale covering the grade levels will be necessary to make across grade level comparisons. Joe Willhoft and Yoonsun Lee of OSPI discussed scaling and equating at the WERA Winter Conference (2007).

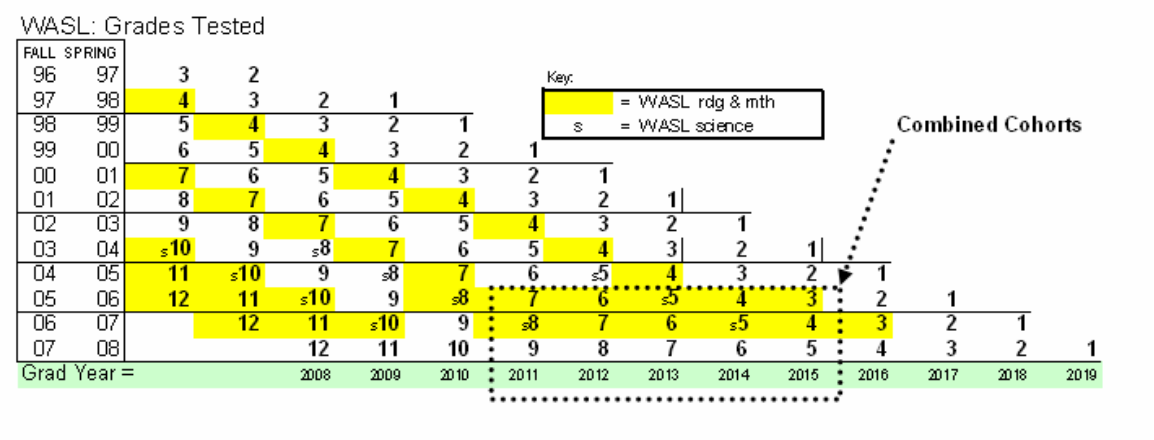


Table 1. Highlighting shows adjacent grades tests.

Comparing groups of students within a grade level is also problematic. School buildings and classrooms have different compositions of students as to achievement levels, socio-economic status, special programs, etc. Comparing an honors class to a remedial class in terms of WASL growth is not appropriate or accurate.

If students were randomly distributed in classrooms and buildings and took the same pre and posttests we could use the average change of WASL scaled score points from 2006 to 2007 within a grade level to compare growth. However, students are not randomly distributed, so we must use techniques to control for different tests and different variables.

It is somewhat intuitive to simply subtract the 2006 WASL result from the 2007 WASL results to measure growth for individual students and then average these differences for groups of students. While this method should not be used for comparing groups of students, it can be useful in beginning to understand a WASL growth model. It is also easily understood and can be readily calculated by anyone with access to WASL results. I would advocate that WASL scores should be made available to teachers and principals in a format that facilitates WASL growth calculations.

Control for Differences

The difference between 2007 WASL Reading scores and 2006 scores for students in grades 4 to 8 was an average of +0.3 points with a range of -81 to +91. In math the range is from -119 to +96 with an average of 0.4 point improvement.

WASL reading (grades 4 – 8) in classrooms with greater than 10 varied from +13 scaled score points to -19 points (a range of 32 points). Math Difference scores varied from +26 to - 29 points. These results reflect students assigned to elementary teachers or to subject area middle school teachers. They do not take into account mobility, team teaching or sharing students between teachers.

Since the WASL is a different test for each grade level one can begin to equalize these differences between grade levels by subtracting the mean difference score for the district for each grade level from every student's score so that one can compare one grade level to another. While this does not address many other variables, it improves comparability between groups.

(Continued on next page)...

When one begins to analyze average growth adjusted for the district mean one notices that level 1 students (based on the WASL pretest) grow the most while level 4 students grow the least with level 2 and level 3 in a middle range of growth. This regression toward the mean is observed at nearly all grade levels. In an attempt to control for differences between achievement levels of students, the averages for each level of students on the pretest could be used to adjust each student's growth scores (subtract the average difference for each WASL level on the pretest from each students growth score).

Another method on investigating WASL growth would be to convert all scores on the pretest and posttest to standard scores to normalize the distribution for each grade level and subject and then calculate the difference between pre and post standard scores (Sattler, 2008). Scores can be converted to z scores by subtracting the mean from each students score and dividing by the standard deviation for each grade and subject. Ideally, the mean and standard deviation would be used from the population of the State results. Lacking access to that data I used the means and standard deviations of my cohort group (approximately 900 students per grade level).

I described four methods of looking at WASL growth. The methods increase in difficulty of calculation and in accuracy or fairness in comparing groups.

1. Difference Scores (post-test minus pre-test),
2. Difference Scores adjusted by district means,
3. Difference Scores adjusted by district WASL levels' means and
4. Standard Z score differences.

Figure 2 displays the first method, Difference Scores, applied to math results over grade levels (grades 4 – 8). For example, students in 6th grade on average lose 9 points versus a gain of 10 points for students in 7th grade. Each year is a different test.

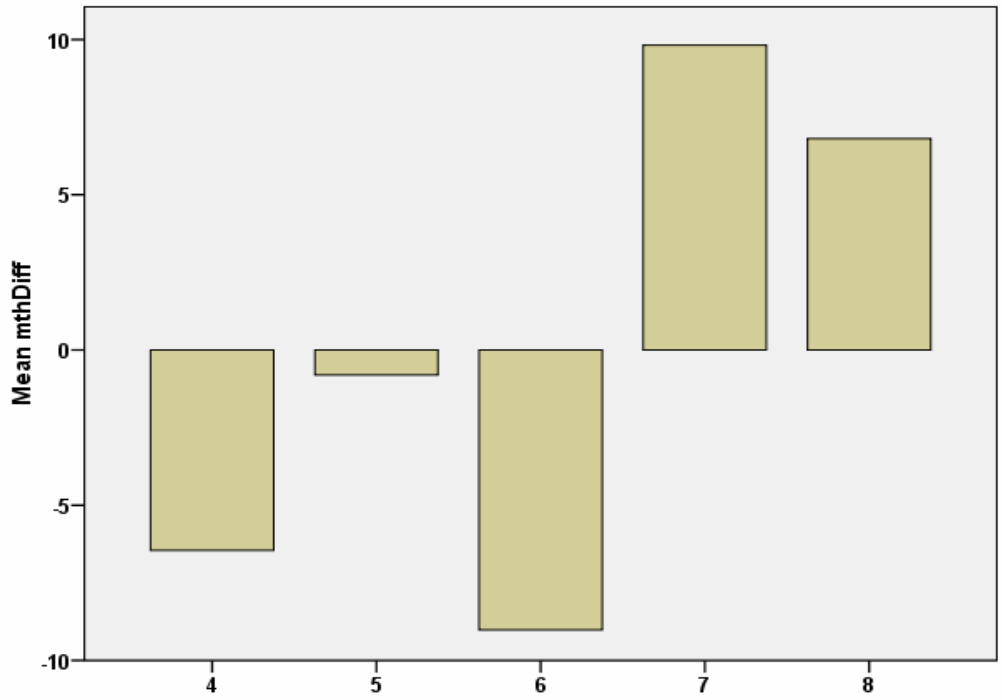


Figure 2. Difference scores, grades 4 to 8

(Continued on next page)...

Methods 2 and 3 adjust scores to the mean resulting in no differences between grade levels in scores for the district. Subgroups of the district can be compared more fairly using these methods.

Each method was applied to math scores for combined grades for each school. The Mean Math Difference indicates elementary schools vary in average math growth from minus 18 points to plus 7 points for grades 4 and 5. Middle Schools display a range of + or – 11 points in math for grades 6, 7 and 8. When scores are adjusted for the district mean the range narrows as it does for adjusting by levels, particularly for middle schools.

Figures 3 – 6 display average growth in math by the four methods in order of presentation for buildings across grade levels (grades 4 – 8). Buildings are in order of least to most Free & Reduced percentages. The charts are illustrative on the effects of the four methods of calculating growth by buildings. Methods 2, 3 and 4 result in strikingly similar profiles.

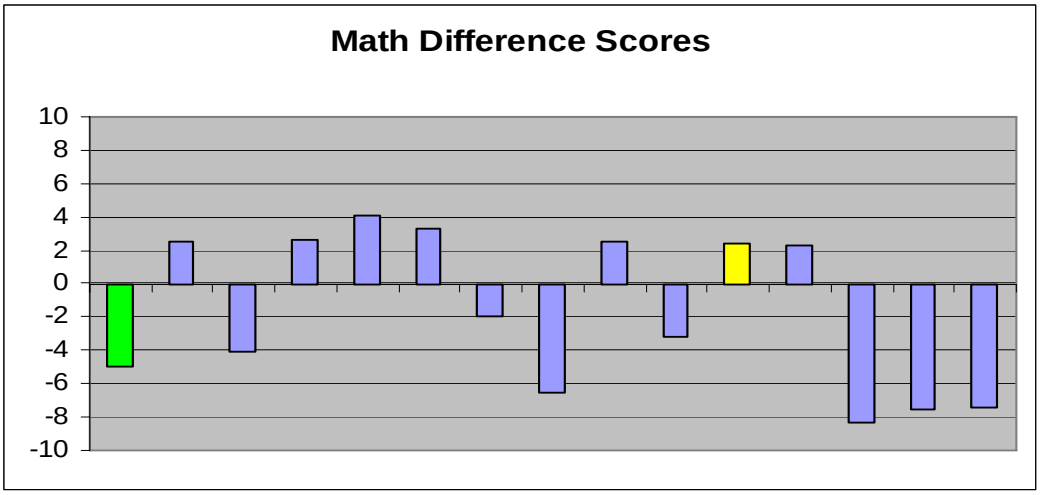


Figure 3. Difference scores, high to low poverty, (left to right).

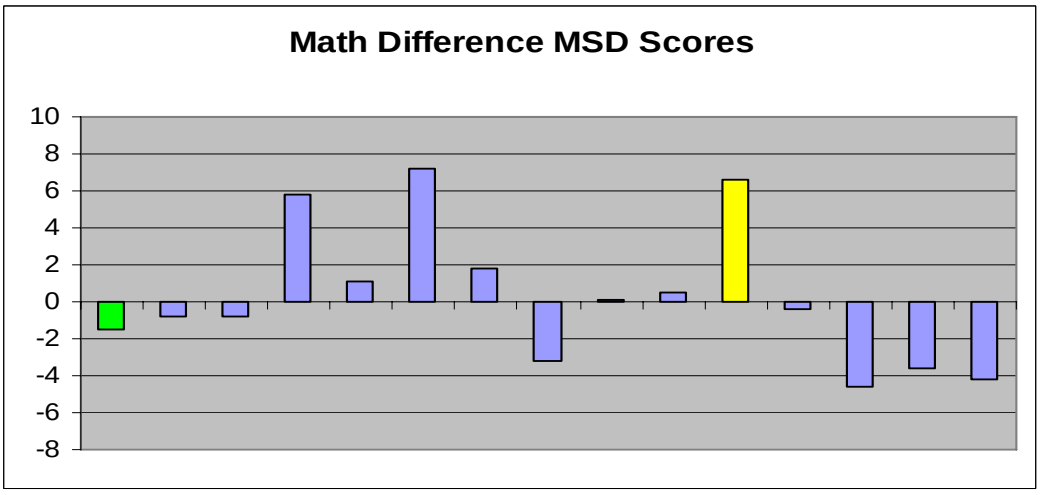


Figure 4. Difference scores, high to low poverty (left to right) adjusted for district mean.

(Continued on next page)...

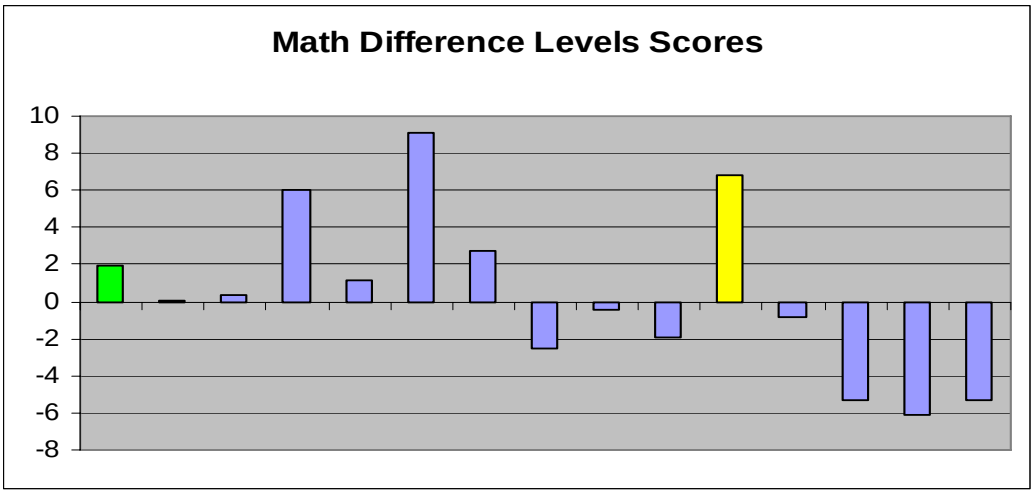


Figure 5. Difference scores, high to low poverty (left to right) adjusted by levels.

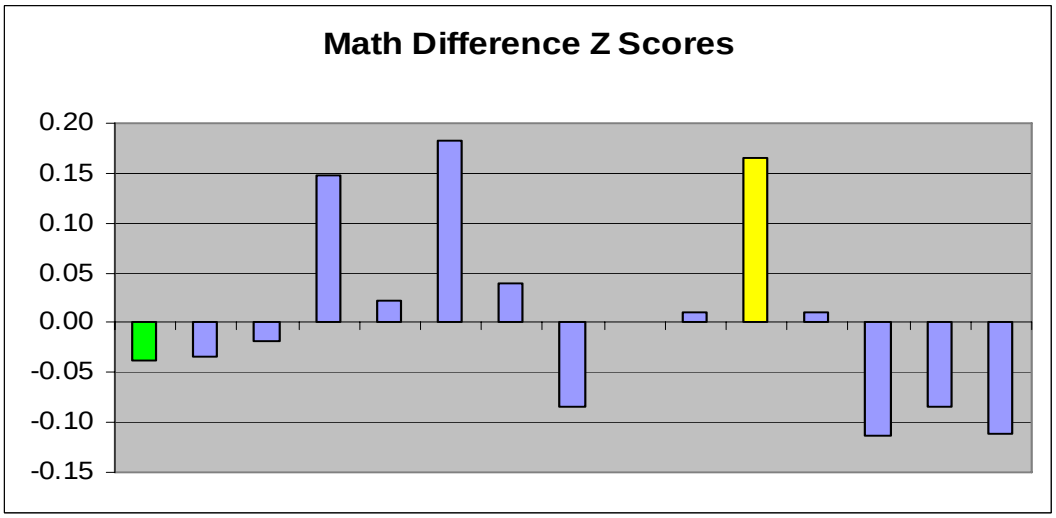


Figure 6. School with highest percentage meeting standard in green (lowest poverty level). Yellow school, highest performing Title 1 School.

The school in green (first bar on left) is a top performing school in the district in terms of percent meeting standard on the WASL. It has a higher proportion of students in level 3 and 4 than other schools. These students do not make the same growth in WASL scaled scores as lower performing students. This is very apparent in figure 7, difference scores without adjustments, but WASL growth is increased by the other methods.

The school in yellow (fifth bar from right) is the highest performing school in terms of growth among Title 1 eligible schools and close to the highest for all schools. Once again methods 2 through 4 yield very similar results.

Discussion:

All methods for analyzing WASL growth have limitations. None of them control for all variables that affect WASL growth and variables in the tests. Consequently, they are not useful in comparing groups of students on a rank order basis. They can, however, be useful in identifying, for instance, the top ten percent of groups as compared to the bottom ten percent to assist in identifying real differences in achievement and promising instructional practices.

(Continued on next page)...

Adjusting difference scores by subtracting the mean, method #2, yielded very similar results to converting all scores to z scores and calculating the differences. It is a fairly simple calculation but should not be used for rank order comparisons.

Z scores calculations should increase in accuracy by using state means and standard deviations when they become available. These methods do not fully control for differences in the WASL tests. A vertical scale is being developed at the state level and is planned to be available this spring for reading with math available at a later date. This common scale should greatly increase our ability to accurately compare WASL growth scores and to more accurately answer the question, "How much on average, did the **same** students' performance change?" Additional information and detail on the four methods of calculating growth can be found on a 2008 Spring WERA presentation on WASL Growth at: http://www.weraweb.org/pages/activities/WERA_winter07/Workshop%205.ppt

References:

Goldschmidt, Pete and Choi, Kilchan. (2005).Policymakers' Guide to Growth Models for School Accountability: How Do Accountability Models Differ? Washington, DC: Council of Chief State School Officers.

Sattler, Jerome M. (2008). Assessment of Children: Cognitive Foundations. San Diego, CA: Jerome M. Sattler, Publisher, Inc..

Schmitz, Don (March 27, 2007). WASL Growth: A Model for Program Exploration. Retrieved April 30, 2008, from Washington Educational Research Association Web site: http://www.wera-web.org/pages/activities/WERA_spring08/Session%201-8.pps

Willhoft, Joseph & Lee, Yoonsun (Dec. 5, 2007). Scaling and Equating. Retrieved April 30, 2008, from Washington Educational Research Association Web site: http://www.wera-web.org/pages/activities/WERA_winter07/Workshop%205.ppt

-Don Schmitz is the Director of Assessment and Program Evaluation for Mukilteo Schools. He is a frequent presenter at WERA conferences with a particular interest in data display and analysis. Contact: SchmitzDJ@Mukilteo.wednet.edu

Everything is on the Web: a Brief Review of Current Research Techniques

By Kate Corby, with Laura Lillard

Anyone who graduated from college more than 5 years ago and hasn't done a lot of literature searching since would likely feel lost in today's search environment. Tools and techniques are changing rapidly and while Google and Wikipedia sometimes feel like the answer to every question, for many researchers, they also engender a sense of discomfort. As more information has become available on the Internet, picking the best sources from amid search results has gotten harder; libraries are seeing an upswing in the number of patrons seeking assistance. It is definitely possible to "do it yourself" but researchers who take the time to update their skills will reap benefits of confidence and timesavings. Here's a brief look at the current landscape for educational researchers.

Access to information: Open Access has gotten a lot of press lately. Congress has actually passed a law requiring that reports of certain types of research conducted with federal funds be freely available to all. So far, these provisions focus on the science, technology and medical fields. For educational researchers, most research reports are still published in journals that require a subscription or download fee for access. In the past researchers who needed something not readily available used an informal network to get copies of articles from friends at a university or by request directly from the author. Publisher copyright agreements are changing and many now allow authors to post the texts of their papers on their own web pages or in an institutional repository. The faculty at Harvard recently agreed to post all of their publications in an open repository. Use any of the following to search for free online versions of subscription materials:

- Google Scholar http://scholar.google.com/
- OAIster http://www.oaister.org/
- Intute http://irs.ukoln.ac.uk/
- Open Doar http://www.opendoar.org/

In the state of Washington, many public libraries have subscriptions to aggregations of journal and magazine content. While there may be little of interest to the educational researcher on the shelves of a public library, these aggregator products often include a rich mix of scholarly and mass-market materials. Usually a numbered library card is all that is needed for remote access to this content. Visit http://www.libraries.wa.gov/ for a link to a nearby library, to see what specific libraries have available.

Google Scholar is one resource that deserves special mention. At one level it is "just" Google but the Scholar search pulls out all the research material from a general Google search. The results are much less junk to wade through and a familiar interface that presents recognizable citations, even on the first summary results page, definitely speeds things up. I use it for known item searches, i.e. looking for things when I know some of the details (usually title or author). It is less helpful for subject based searching. Figure 1 shows the beginning results of a typical subject oriented Google Scholar search. While the page is attractive, the tendency of the Google search algorithm to pull up older classic materials is often not useful for researchers. They keep improving it however; notice the "Recent articles" link at the top of the results page, and the "Key authors" list at the bottom. See below for additional, more effective, ways to search by subject/keyword. Note the "cited by" line below the citation in Google Scholar, it is a link to more current material that has cited these articles. That feature can be a huge help in updating from an older article that has, perhaps, served as key in an earlier related project.

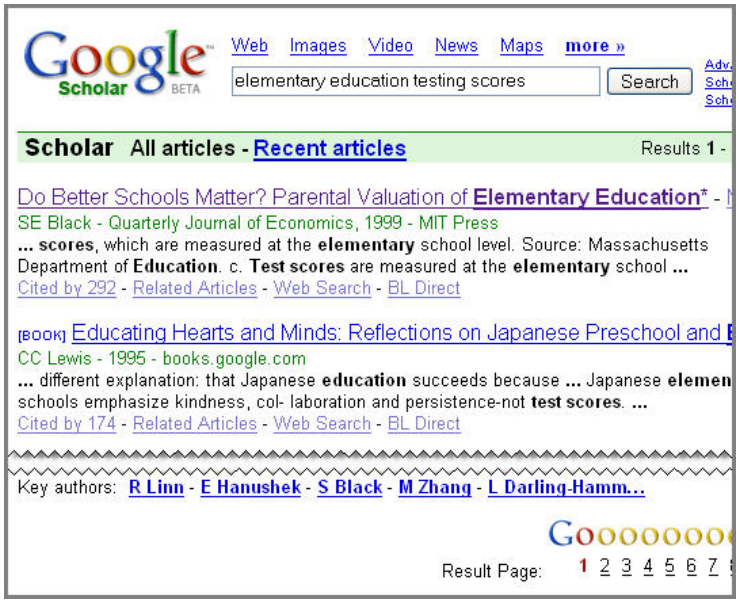


Figure 1. *Google Scholar results page.*

(Continued on next page)...

Searching for the unknown. Looking for information on a specific topic makes Internet searching seem particularly unfriendly. This is when searches usually yield large numbers of results and the searcher feels unsure of finding the best and most recent information on the topic. Often experienced researchers try to overcome this problem by searching the contents of just one or two known journals. Extending the logic of such a strategy to search a limited but slightly larger universe is probably a better idea. There are so many new sources becoming available – tried and true journals might not be the best bet. ERIC has a new web interface that works well for subject searches. http://www.eric.ed.gov/. Try the "Advanced Search" page to get a form with enough windows for a complex search. See figure 2. Notice that searchers can look up and use uniform descriptors to increase the relevance of results. There are several other ways of limiting results (years, publication type, education level) Be careful about checking the box to show only free full-text from ERIC, since that will effectively eliminate journal articles from your results. As mentioned above, there are other sources of free full text journal articles, once a citation is in hand.

The ERIC feature "My ERIC" allows users to save searches and create folders of results for future use.

In a new trend having an impact on research, journal publishers are making their complete line of journals searchable on their branded web sites. This trend it problematic because it limits results to a particular publisher which limits search results artificially and can skew that publisher's impact on future research. It is important to keep this issue in mind; one should never use such a search as a basis for new research. Still for finding a quick update on a topic, many researchers find these tools helpful. The American Educational Research Association recently placed all of its journals with Sage. The Sage interface is online at http://online.sagepub.com/. Once again there is an advanced search screen that allows more complex searches. Notice the facility, under "My Tools" – see figure 4 -- to register for a free searcher's account that will allow anyone to save searches and direct alerts to an email box whenever new articles of interest are published.

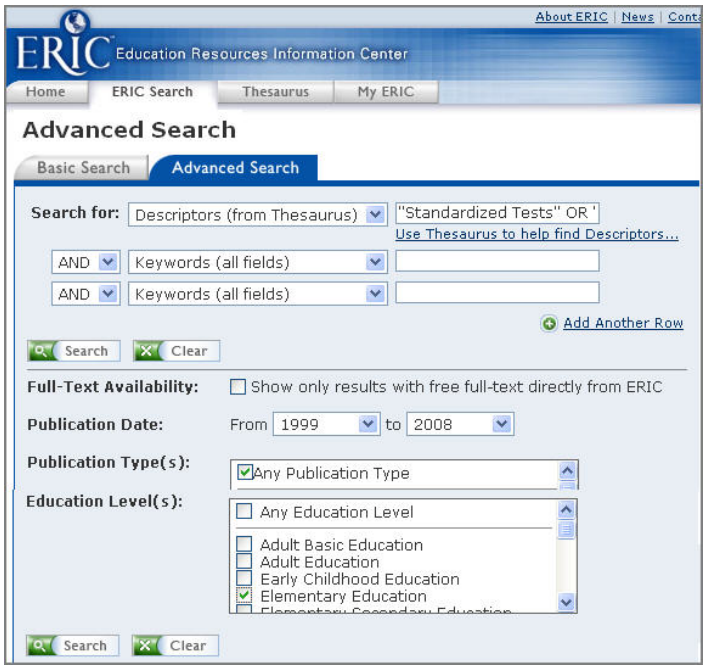


Figure 2. ERIC Advanced Search.

In figure 3, notice that once ERIC results come up it is possible to sort so that the most recent come up on top.

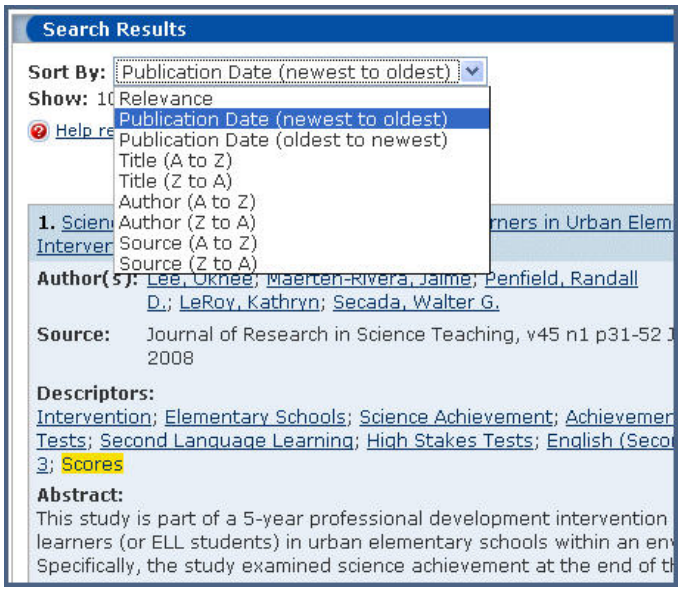


Figure 3. ERIC Search Results page.

(Continued on next page)...

Figure 4. *Sage-My Tools-Alerts sign up.*

It is true that the task of literature searching is changing rapidly, but it is changing toward ease and ubiquity. A few minutes of extra time finding the right place to search and doing a precise search can save hours of wading through long lists of irrelevant citations, or missing important articles.

-Kate Corby is Education and Psychology Librarian at Michigan State University, East Lansing, Michigan.
http://www.lib.msu.edu/corby/

-Laura Lillard is Education Librarian at the University of Washington, Seattle.
http://www.lib.washington.edu/subject/librarians/lillard.html

Using Students' Primary Languages to Support Reading in a Multilingual ELL Classroom: Bilingual Concept Mapping to Facilitate Critical Thinking and Vocabulary Development

By Richard Bailey

Abstract

Concept maps have shown great promise as tools to guide learners through complex thought processes; however, very little is known about the effective use of bilingual concept maps with ELL students. This action research project examines the use of bilingual concept maps in vocabulary, inference, and main ideas and supporting details among beginning and intermediate level English language learners (ELLs). It was found that students scored higher on teacher made assessments after completing bilingual concept maps on vocabulary and inferencing tasks, but not on identifying main ideas and supporting details in reading. Furthermore, issues of language transfer and primary language loss hindered the performance of students at higher levels of proficiency in English on bilingual mapping tasks. The findings of this study are examined in terms of cognitive processing models and the language threshold needed in order to receive positive cognitive benefits from interactions between first and second languages, and recommendations are given for further research and practice in supporting English instruction with primary language use.

An ongoing dilemma among programs for English Language Learning (ELL) in Washington State centers on how to use students' primary or first languages when helping them learn English. Those who advocate bilingual education will point to studies such as Thomas and Collier (Collier & Thomas, 2004; , 1997) who found that students in bilingual education programs tended to outscore students in all English programs on tests given in both the students' primary language and in English. On a national level, bilingual education advocates also urge the use of bilingual methods in order to show that schools value and respects the linguistic and cultural backgrounds of their students (Au, 1998; Jim Cummins et al., 2005; Jim Cummins, Chow, Schecter, & Yeager, 2006) while also building on the "ways of knowing" or problem solving skills and bodies of knowledge already developed within the home and community (Moll, Amanti, Neff, & Gonzales, Spring 1992; Stagg Peterson & Heywood, 2007). While proponents of all English ELL programs certainly acknowledge the value of learning in two languages, they question the feasibility of using students' primary languages in a multilingual classroom. They argue that bilingual education just is not possible when so many different languages are present. They also point to the question of fairness by saying that it is inherently unfair to offer bilingual education to one group of students while another group, perhaps a language group represented by a very small number of students, goes unserved due to limited instructional resources in their primary language. Acknowledging the valid points of both sides of this argument, this study seeks to understand how students' primary languages may be used to help facilitate learning even in a multilingual setting. As an ELL teacher in a multilingual, content-based program, I used my students' primary languages in concept maps to target specific instructional goals in vocabulary, inference, and discerning main ideas and supporting details in reading.

In designing the study, I drew on research from second language reading, concept mapping, and from research on cognitive processes in bilingual individuals. My goal was to find an intersection between current theory on how bilingual students think and how I might use those thinking processes to best tailor my instruction to their processing abilities. In looking for a cognitive processing model for bilingual thinking, I drew primarily on the revised hierarchical model espoused by Kroll and Stewart (1994).

They argue that conceptual representations of words are held in memory independently of language, but they are more readily accessible when using primary language (L1). Furthermore, they argue that even bilingual individuals who are highly proficient in both languages often have a larger lexical store in their L1 than they do in their second language (L2). For this reason, cognitive connections in thinking from L2 to L1 are often made at the lexical level.

(Continued on next page)...

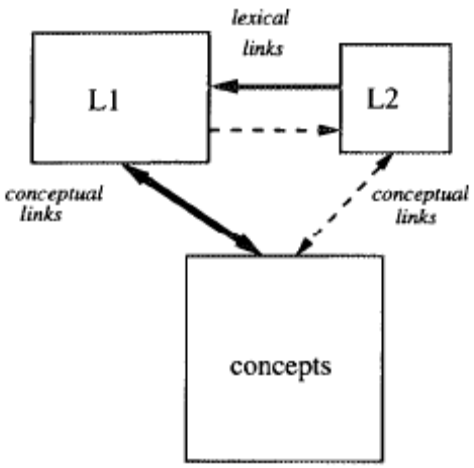


FIG. 3. Revised hierarchical model of lexical and conceptual representation in bilingual memory.

(Figure 1) (Kroll & Stewart, 1994 p. 158)

Individuals will then retrieve conceptual information through their L1 and then translate back into L2 in order to communicate those understandings. According to the model, it actually takes more time for bilinguals to translate from L1 to L2 than from L2 to L1 because of the strong lexical store in L1. They hold multiple words for the same object and multiple meanings for the same word.

The revised hierarchical model helps explain how bilingual students process vocabulary tasks, but because Kroll and Stewart limit the tasks in their study to picture recognition and word naming, their study does not completely explain how individuals process higher level thinking skills needed for inference making and discerning main ideas and details from written text. For that information, I used the work of Ruddell and Boyle (1989) who found that concept mapping appears to reduce memory demands therefore allowing students to focus their attention on "the integration of details to support main ideas in their written expository summarizations" (p. 21). In other words, the use of concept maps helps the learner to organize thoughts on paper. Once those thoughts are on paper, memory processing can take on other functions such as engaging in procedural tasks related to forming conceptual understandings.

The Ruddell and Boyle study used monolingual English speakers and focused on summarization tasks alone. More recently, however, concept mapping has been applied to thinking skills in second language learning by popular approaches such as the Cognitive Academic Language Learning Approach (CALLA) (Chamot & O'Malley, 1994; O'Malley & Uhl Chamot, 1990) and the strategy instruction advocated for in the Sheltered Instruction Observation Protocol (SIOP) (Echevarria, Vogt, & Short, 2004). However, in my review of the literature, I found no study that used a participant's primary language in concept mapping as an aid to improving cognitive processing in L2. It seems that bilingual students are being asked to take on extra cognitive processing demands for both vocabulary and conceptual tasks. Therefore, my study sought to ease this processing demand by integrating students' primary languages into concept mapping tasks. I theorized that by allowing students to support their ideas by using their L1, they would reduce cognitive demands on memory and allow them more "mental space" to form lexical links between L1 and L2 for vocabulary words. In terms of developing critical thinking skills, I theorized that the bilingual concept maps would again free up valuable mental space that would then be used for forming inferences and main ideas.

Study Description

The study was conducted as part of a professional growth plan using existing curriculum which included a series of readings from the Visions textbook series(McCloskey & Stack, 2004a, 2004b; O'Sullivan & Newman, 2006). A design-based or action research model was used. This model advocates the use of multiple iterations of key instructional methods in order to better understand the learning that occurs around the different implementations and the changes that need to be made to the instruction in order to better assist students in the learning process(Design-Based Research Collective, 2003; Shavelson, Phillips, & Feuer, 2003). Students read stories and expository text from the appropriate level of the Visions textbook. Stories were taught in two different categories. The first category was inference, and the second category was main ideas and supporting details. Before each story, students received pre-teaching in background knowledge, and they were given a vocabulary concept map. The first iteration in each category was conducted all in English. Subsequent iterations in each category were conducted using bilingual concept maps for vocabulary in all stories and for inference and main ideas and supporting details in the respective categories.

(Continued on next page)...

During the reading of each story, students participated in teacher led discussions about key vocabulary and ideas. Students were asked to predict, ask questions, and make inferences. Concept maps were given out during the reading of the story, and students were asked to complete them while they were reading. Students were assessed on both vocabulary and conceptual understandings in inference and main ideas and supporting details.

Participants

The study was conduced in three of my high school level ELL classes. The students were enrolled in the classes according to their ELL proficiency level as assessed by the Washington Language Proficiency Test (WLPT II). Students were in grades nine through twelve, and they ranged in age from fourteen to eighteen years old (see table 1).

Beginning Level	Number of students	Total Male	Total Female
Primary Language			
Spanish	1	1	
Korean	1	1	
Olmeric	1	1	
Vietnamese	1		1
Russian/Ukrainian	3	2	1
Total	7	5	2

(Table 1: Beginning Level Language and Gender)

Intermediate Level	Number of students*	Total Male	Total Female
Primary Language			
Spanish	5	2	3
Russian/Ukrainian	2	1	1
Cambodian	1		1
Tagalog	1	1	
Tai	1		1
Chinese	1	1	
Total	11	5	6

***Three students were eliminated due to incomplete test data. Their data is not included here.**

(Table 2: Intermediate Level: Language and Gender)

The beginning level students were students who had scored a level 1 on the WLPT II. These students were placed in a beginning English classroom where they were receiving instruction from the introductory level of the Visions textbook series(O'Sullivan & Newman, 2006). As the study progressed, this group transitioned from the introductory level to intermediate level A(McCloskey & Stack, 2004a) in the series. Stories used in the study came from both levels of the Visions series.

The second group was formed of students who scored at the intermediate level, level 2, on the WLPT II exam. These students were placed in a beginning English classroom where they were receiving instruction from the introductory

(Continued on next page)...

level of the Visions textbook series(O'Sullivan & Newman, 2006). As the study progressed, this group transitioned from the introductory level to intermediate level A(McCloskey & Stack, 2004a) in the series. Stories used in the study came from both levels of the Visions series.

The second group was formed of students who scored at the intermediate level, level 2, on the WLPT II exam. This group was made up of two different classes. One class was taught by me. The other class was taught by a paraeducator who used materials and instructional methods developed by me. This paraeducator is a former science teacher who holds a teaching certificate in physics and an endorsement in ELL. The intermediate group read stories and expository articles from the Visions textbook series level A. As the study progressed, they transitioned from the intermediate level A to the more advanced level B(McCloskey & Stack, 2004b). Stories involved in the study came from both books.

A third group of students who scored at the advanced level, level 3, on the WLPT II originally began the study, but was eliminated due to ethical considerations which will be discussed later.

Data Collection and Analysis

Data collection included assessment data as well as other qualitative information such as notes on student discussions, classwork, concept maps, and teacher observations. This data was analyzed for trends in student work. Those trends were considered in light of current education theory on primary language use in all English contexts and the development of thinking skills in bilingual learners.

Students' vocabulary understanding was assessed using a fill in the blank approach which is explained by Hughes (2003) as a highly reliable approach that offers the instructor enough control over test items that questions can be constructed in such a way as to make it unlikely that students will make wrong choices based on misunderstandings of the passages. Learning goals for inference and main idea and supporting details were assessed using short answer and extended response questions questions. Again, assessment approaches that according to Hughes(2003) and Brown (2004) are appropriate and effective assessment practices for higher level thinking objectives. It was also intended that these test items mirror WASL questions that students would encounter later in the year. Responses were scored using a rubric that translated into letter grades. No per cent score was given to students. For the purpose of this study, however, a point system was assigned to each letter grade based on a percentage basis. See table three for the results of these tests.

Test Data

The intention of this study was to understand how primary language can support thinking skills in a classroom environment. For this reason, random assignment was not possible, and the number of participants is too low to justify further statistical analysis. It would be inappropriate to draw generalizable conclusions from the data in table 3; however, it is possible to identify trends that may warrant investigation in further research. Four trends will be discussed here. The first two involve positive trends in both vocabulary and inference scores when bilingual concept maps were used. The third involves an explanation as to why that same trend was not observed when students used bilingual concept maps for identifying main ideas and supporting details. The fourth trend comes from student conversations and my observed resistance of students at higher proficiency levels to engage in bilingual concept mapping tasks.

Trends in Vocabulary

In Category 1 tests, students at both proficiency levels demonstrated a noticeable increase in test scores on vocabulary when bilingual concept maps were used. The same increase does not happen in Category 2 tests. After seeing the initial increase in Category 1 scores, I wondered if the dip in scores on the Category 2 Vocabulary (1) test was due to poor instruction, the ineffectiveness of bilingual concept maps, or the fact that the test was given on the day before spring break, and student motivation may have been an issue. I decided that the decline in scores may have been due to a long wait time (almost two weeks) between the completion of the bilingual concept map and the actual test. It may be possible that students need continued practice using the words in both languages before they are prepared to take the test.

(Continued on next page)...

Category 1: Inference

All English	**Beginning Level**	N=7	
		Mean	Standard Deviation
Vocabulary		6.43	1.81
Inference		6.86	1.46
Bilingual			
Vocabulary		7.57	3.36
Inference		8.29	0.71

Category 2: Main Ideas and Supporting Details

All English	**Beginning Level**	N=7	
Vocabulary		7.29	1.8
Main Ideas and Supporting Details		7	2.31
Bilingual			
Vocabulary (1)		5.86	3.02
Vocabulary (2)		6	1.63
Main Ideas and Supporting Details (1)		7.29	2.14
Main Ideas and Supporting Details (2)		7.29	1.6

Category 1: Inference

All English	**Intermediate Level**	N=13	
		Mean	Standard Deviation
Vocabulary		4.14	2.25
Inference		4.5	1.74
Bilingual			
Vocabulary (1)		5.33	3.55
Vocabulary (2)		7.7	2.99
Inference (1)		5.77	2.7
Inference (2)		7.7	1.35

Category 2: Main Ideas and Supporting Details

All English	**Intermediate Level**	N=13	
		Mean	Standard Deviation
Vocabulary		5.72	3.17
Main Ideas/Details		7.55	1.92
Bilingual			
Vocabulary (1)		4.36	2.9
Vocabulary (2)		6.85	2.41
Main Ideas/Details (1)		5.29	3.11
Main Ideas/Details (2)		6.15	2.64

(Table 3: vocabulary and concept scores for beginning and intermediate students)

(Continued on next page)...

For the next reading (Vocabulary (2) test scores), I changed the instructional approach. After the initial introduction of vocabulary words and the use of the bilingual concept maps, students on subsequent days were asked to complete short writing assignments that used the vocabulary words in both their L1 and L2. The test was given upon completion of the story reading, continued discussion, and the completion of three different bilingual writing assignments. Scores on Vocabulary (2) tests in Category 2 do show an increase over the Vocabulary (1) scores. At the beginning level, the increase in Vocabulary (2) scores was still not enough to overcome the 7.29 mean achieved on the all English story used in Category 2. At the intermediate level, the increase in scores for the Vocabulary (2) test was more than the Vocabulary (1) and the all English vocabulary test in Category 2. The data may indicate an overall higher degree of success when bilingual concepts are used in conjunction with daily bilingual practice of vocabulary words. In accordance with the revised hierarchical model, it may take longer than expected to help students develop and maintain lexical connections between words in L2 and their links to the appropriate words in L1.

Trends in Inference

The largest improvement between all English concept maps and bilingual concept maps was seen on inference tasks. Students in both categories and in all iterations improved their performance when bilingual concept maps were used. I attribute these findings to an increased efficiency in processing brought on by the bilingual concept maps. Other authors have pointed out that inference is an area that receives positive transfer from L1 to L2 (Grabe & Stoller, 2002; Pulido, 2007) on reading tasks both at the conceptual level and at the lexical level. It may be that the use of the bilingual concept maps aided in the formulation of stronger links between lexical stores in L1 and L2 and between those respective lexical stores and conceptual information held in long term memory. The stronger connections may have eased processing capacity thereby freeing up students to engage in critical thinking about inference passages in the story. Hulstijn et al (2007) explain this ease in processing capacity as the processing efficiency hypothesis which states that "when lower order processes are slow and attention demanding, the higher order processing needed for text comprehension suffers" (p. 478). It could be that the use of the students' primary languages eased processing capacity thereby allowing them to concentrate on inference tasks. If so, this would extend Hulstijn el al's definition of the processing efficiency hypothesis to include inference skills.

Trends in Main Ideas

The data showed an overall negative trend in the use of bilingual concept maps to support the identification of main ideas and supporting details. For students at the beginning level, scores were essentially the same between iterations that used all English concept maps and those that used bilingual ones. At the intermediate level, scores on the iterations using bilingual concept maps were noticeably lower. Grabe and Stoller (2002) point out that linguistic and cognitive skills in phonological awareness, topical knowledge, general background knowledge, problem-solving strategies, and inferencing skills can all transfer from L1 to L2 (p. 54), but they do not always transfer in a positive fashion. It may be that the identification of main ideas and supporting details is one of those skills that should be added to the list of abilities that do not necessarily transfer positively.

Another possibility is that the use of bilingual concept maps may have actually hindered student performance in identifying main ideas and supporting details. According to the revised hierarchical model, lexical links are strongest when moving from L2 to L1, and conceptual links are strongest between L1 and the conceptual store in long term memory. The intention of the bilingual concept map was to help facilitate this process by guiding students as they translated L2 text into L1 where they could use reasoning skills to determine if the information was a main idea or a supporting detail. What students may have been doing instead was to bypass L1 and work directly between L2 and their conceptual store in long term memory. The extra translation I was asking of them may have hindered the connection they were trying to make between their L2 lexical store and long term memory.

(Continued on next page)...

Since their lexical store in L1 is greater than in L2, the translation may have actually acted as a hindrance by causing them to look for extra words and word meanings that fit the context of the reading. If further research were to substantiate this finding, it may be necessary to refine the revised hierarchical model to indicate which cognitive functions make strong and weak connections between which lexical store.

Students Eliminated from the Study

A final note on observed trends in this project has to do with those students who had difficulty with this work. Many students, especially those with higher levels of proficiency in English expressed frustration and concern with the bilingual assignments. This was especially true for a group of students who scored at the advanced level of proficiency on the WLPT II, level 3. These students were originally supposed to be included in the study, but they were dropped after it became clear to me that many of them did not have the linguistic resource in L1 to successfully complete the activities. This was also true for several of the students at the intermediate level. When asked to complete bilingual concept maps in vocabulary, they replied that they did not know the words in their first language. Students were told to look the words up in their bilingual dictionaries, but once they found the words in the dictionaries, they still did not indicate that they had had any previous exposure to the words. These students were encountering new words in both English and in their first language.

The conversations students had with me regarding their lack of confidence in the academic vocabulary and conceptual understandings available to them in their L1 reminded me of a well known but perhaps now underused term referred to as subtractive bilingualism. Originally described by Lambert in 1974 (as cited in Baker, 1993), the term has come to be used in two different ways. First, it refers to a distinction between additive and subtractive features of language and cultural development. When a language is acquired with little or no pressure to "replace or reduce" (p. 95) the first language, then the process of language acquisition is thought of as additive. When the second language is acquired in a way that subordinates or even eliminates the first language, the process is considered to be subtractive. Lambert also used subtractive bilingualism to refer to the negative cultural and cognitive benefits of having "underdeveloped" language skills (p. 95). For the students in my classroom, the term underdeveloped appears to mean that they were lacking the vocabulary necessary in their primary language to facilitate academic discussion about literature.

This resistance to bilingual concept maps was not observed with students at the beginning proficiency level. This may have to do with the threshold of language needed to be able to receive cognitive benefits from L1 when acquiring an L2. Cummins(as cited in Baker, 1993) and others have argued that the amount of language available to a student in both L1 and L2 plays an important role in determining the degree to which the student is able to function in an academic environment. Those with limited proficiency in both L1 and L2 will likely be unable to keep up with the class, even in an ELL setting where the pace of instruction is much slower (Baker, 1993). These are the students who we traditionally refer to as limited bilinguals. However, an often overlooked category is those students who have achieved a degree of proficiency in their L2 appropriate for their years of study, but who are still learning their new language. They may not yet be receiving any cognitive benefits such as vocabulary transition and reasoning capabilities from using their primary language as a resource when learning English.

This second group appears to characterize the students in my study who were resistant to the concept mapping activities. By their own reports, these students were at or near grade level when they stopped instruction in their first language and moved to the United States. Once they began learning in English, they progressed comfortably through an ELL program, but there was no more academic instruction in their primary language. Cummins (1977) has stated that the threshold needed among intermediate learners may be different depending on the type of cognitive operations the student engages in and the level of cognitive development the student is at. In other words, it is possible that some of my students were proficient enough in English that they could handle academic tasks

(Continued on next page)...

in English without the support of their primary language. The converse is also true. There may have been some students who due to a lack of vocabulary in their primary language were unable to adequately process the task.

Conclusion, Suggestions for Further Research

In this study, it was observed that students relied on their primary language for support in vocabulary and thinking skills related to their English instruction. These skills were successful to varying degrees depending on proficiency level and the cognitive task at hand. This observation is consistent with other theories about language threshold and transfer of L1 skills (Baker, 1993; J Cummins, 1977; Grabe & Stoller, 2002). The amount of language needed to undertake academic tasks in L2 varies depending on the cognitive demands of the task. Further research should seek to flush out which conceptual tasks require what degree and type of language skills in L1 and in L2 in order to make for a positive and useful transfer.

Finally, further investigation into this issue should seek to understand ways to combat the negative effects of subtractive bilingualism. One possible strategy for enhancing students' primary language skills may be to incorporate primary language activities into the all English school culture and curriculum. In this study, higher scores on vocabulary and inference tests were obtained when bilingual concept maps were used. It remains to be seen if other thinking skills such as cause and effect or making generalizations can benefit from the same type of mapping activity. It may also be possible that bilingual concept mapping, even in an all-English, content-based program, that starts with lower proficiency levels and continues on through to the more advanced level may help students maintain their primary language, thereby reducing the resistance that was observed in this study by students at the higher proficiency level. Further research should seek to understand if the integration of primary language into other aspects of all English instruction would increase student learning in other areas. For example, bilingual writing assignments, heritage language clubs, or even bilingual presentations may work to preserve a student's primary language while also increasing the cognitive benefits in the learning of English.

Works Cited

Au, K. H. (1998). Social Constructivism and the School Literacy Learning of Students of Diverse Backgrounds. Journal of Literacy Research, 30(2), 297-329.

Baker, C. (1993). Foundations of Bilingual Education and Bilingualism. Clevedon, England: Multilingual Matters.

Brown, H. D. (2004). Language Assessment: Principles and Classroom Practices. White Plains, NY: Longman.

Chamot, A. U., & O'Malley, J. M. (1994). The CALLA Handbook: Implementing the Cognitive Academic Language Learning Approach. Reading, MA: Addison-Wesley.

Collective, D.-B. R. (2003). Design-Based Research: An Emerging Paradigm for Educational Inquiry. Educational Researcher, 32(1), 5-8.

Collier, V. P., & Thomas, W. (2004). The astounding effectiveness of dual language education for all NABE Journal of Research and Practice, 71(4), 649-655.

Cummins, J. (1977). Cognitive Factors Associated with the Attainment of Intermediate levels of Bilingual Skills. The Modern Language Journal, 61(1/2 (Jan.-Feb.)), 3-12.

Cummins, J., Bismilla, V., Chow, P., Cohen, S., Giampapa, F., Leoni, et al. (2005). Affirming Identity in Multilingual Classrooms: . Education Leadership(September), 38-43.

Cummins, J., Chow, P., Schecter, S. R., & Yeager, B. (2006). Community as Curriculum. Language Arts, 83(4), 297-306.

Echevarria, J., Vogt, M., & Short, D. J. (2004). Making Content Comprehensible for English Learners: The SIOP Model. Boston: Pearson.

Grabe, W., & Stoller, F. (2002). Teaching and Researching Reading. Harlow, UK: Pearson.

Hughes, A. (2003). Testing for Language Teachers. Cambridge: Cambridge University Press.

(Continued on next page)...

Hulstijn, J., van Gelderen, A., Schoonen, R., Stoel, R., & de Glopper, K. (2007). Development of Adolescent Reading Comprehension in Language 1 and Language 2: A Longitudinal Analysis of Constituent Components. Journal of Educational Psychology, 99(3), 477-491.

Kroll, J. F., & Stewart, E. (1994). Category Interface in Translation and Picture Naming: Evidence for Asymmetric Connections between Bilingual Memory Representations. Journal of Memory and Language, 33, 149-174.

McCloskey, M. L., & Stack, L. (2004a). Visions Student Book A. Boston: Thomson Heinle.

McCloskey, M. L., & Stack, L. (2004b). Visions Student Book B. Boston: Thomson Heinle.

Moll, L. C., Amanti, C., Neff, D., & Gonzales, N. (Spring 1992). Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. Theory Into Practice, 31(2), 132-141.

O'Malley, J. M., & Uhl Chamot, A. (1990). Learning Strategies in Second Language Acquisition. Cambridge: Cambridge University Press.

O'Sullivan, J. K., & Newman, C. (2006). Introductory Visions Student Book. Boston: Thomson Heinle.

Pulido, D. (2007). The Effects of Topic Familiarity and Passage Sight Vocabulary on L2 Lexical Inferencing and Retention through Reading Applied Linguistics, 28(1), 66-86.

Ruddell, R. B., & Boyle, O. F. (1989). A study of cognitive mapping as a means to improve summarization and comprehension of expository text. Reading Research and Instruction, 29(1), 12-22.

Shavelson, R. J., Phillips, D. C., & Feuer, M. J. (2003). On the Science of Education Design Studies. Educational Researcher, 32(1), 25-28.

Stagg Peterson, S., & Heywood, D. (2007). Contributions of Families' Linguistic, Social, and Cultural Capital to Minority-Language Children's Literacy: Parents', Teachers', and Principals' Perspectives. The Canadian Modern Language Review, 63(4), 517-538.

Thomas, W., & Collier, V. P. (1997). School Effectiveness for Language Minority Students. Washington, DC: National Clearinghouse for Bilingual Education.

-Richard Bailey is the ELL department head at Cascade High School, Everett Public Schools. He is a doctoral student at the University of Washington. He has written and presented on instructional methods for ELL students and leadership in English language teaching organizations.

A Calculation of Educational Work

By Bruce Dean Kelley, Ph.D.

Despite fourteen years of education reform in the state of Washington, socioeconomic status remains the variable most strongly correlated with institutional success on the Washington Assessment of Student Learning (WASL). This makes comparing the strength of educational programs based upon the percent of students passing the WASL among schools or districts (a favorite pastime of news media and uninformed citizens) a statistical fallacy, unless the schools or districts being compared have very similar demographics.

This study demonstrates a way to statistically remove some socioeconomic effects on middle school WASL scores, allowing more appropriate comparisons of the relative efficacy of educational programming among schools or districts.

METHODS

1. Data used in the first portion of this study were from Washington school districts having total district student populations of 600 or more, to reduce single-teacher effects found in smaller districts. The percent of students passing the Reading, Mathematics, Writing, and Science WASL for grades 6, 7, and 8 were extracted from the Washington State Report Card website, along with district-reported percentages of students qualifying for free or reduced lunches. The mean WASL percentage (gWp) was calculated for each grade level using number of students tested for each grade level and the number of students passing required tests for each grade level. The final mean (mWp) was determined by averaging the means for the three grades. Socioeconomic status (SES) was calculated by subtracting the percentage of students qualifying for free or reduced lunch from one. A scatter plot of SES vs. mWp was drawn (JMP Statistical Software, v. 5.1), with a third-order polynomial fit (Fig. 1).

2. Data for the second portion of the study were used from Washington middle-level schools, to examine differences among individual schools in larger school districts. Schools selected for this study had grade-level populations of 90 or more, to reduce single-teacher effects. Middle-level schools without (at least) grades 6, 7, and 8 were not used. Final mean (mWp) and SES were calculated as above. A scatter plot of SES vs. mWp was drawn (JMP Statistical Software, v. 5.1), with a different third-order polynomial fit (Fig. 2).

3. Following statistical discovery of the district best fit equation, the relative educational work (W) was calculated in the following manner:
    a) For each district, the expected WASL percentage (eWp) was calculated using the district's SES and the best fit equation.
    b) This eWp was subtracted from each district's actual mWp to determine the gap between the expected and actual (difference = educational work = W).

    If W = zero for a district, students performed as expected on the WASL, based on their relative SES. A positive score means that students scored higher than expected from their SES; a negative score means that students scored worse.
4. Similar discovery of educational work values were calculated for <u>schools</u> using the best fit equation from figure 2 and school data.

RESULTS

The Summary of Fit tables shown in fig. 1 and 2 demonstrate that cubic best fit lines are good models for correlation between mean WASL percentages (mWp) and socioeconomic status (SES), with R^2 being higher for schools (0.80) than for districts (0.66), and $p<0.001$ for both analyses. Simple best fit lines also demonstrated significant correlations, but R^2 values were lower (data not shown).

Figures 3 and 4 show that the data for districts and schools were similarly distributed and close to normal. In both cases, the standard deviation was about 0.06.

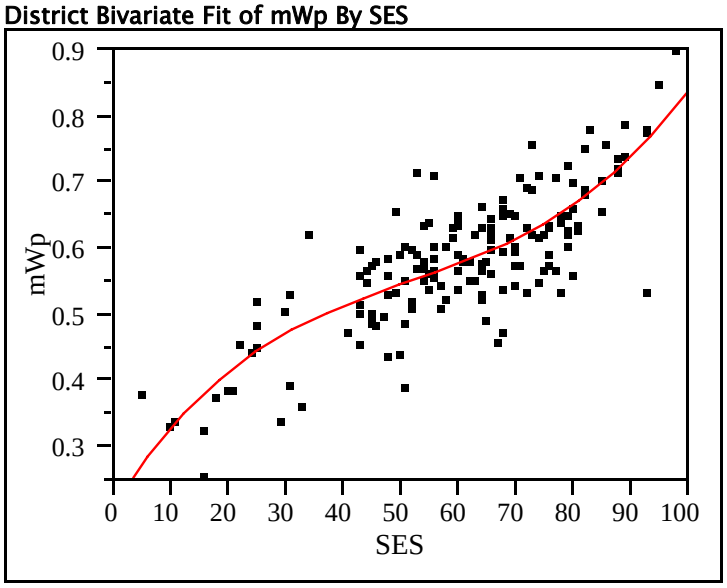
DISCUSSION

The high R^2 values in these two studies show that the socioeconomic status (SES) of the students in a school or district remains a strong predictor of school-wide success on the WASL. Effects of poverty on student achievement test results has been widely demonstrated in the past (Sirin, 2005; Raudenbush, 2004; White, 1982). However, the calculation of Educational Work gives administrators and

(Continued on next page)...

Figure 1.

District Bivariate Fit of mWp By SES



Polynomial Fit Degree=3

mWp = 0.3787975 + 0.0032433 SES + 0.0000277 (SES-59.7317)^2 + 0.0000013 (SES-59.7317)^3

Summary of Fit

RSquare	0.658729
RSquare Adj	0.65233
Root Mean Square Error	0.061615
Mean of Response	0.577543
Observations (or Sum Wgts)	164

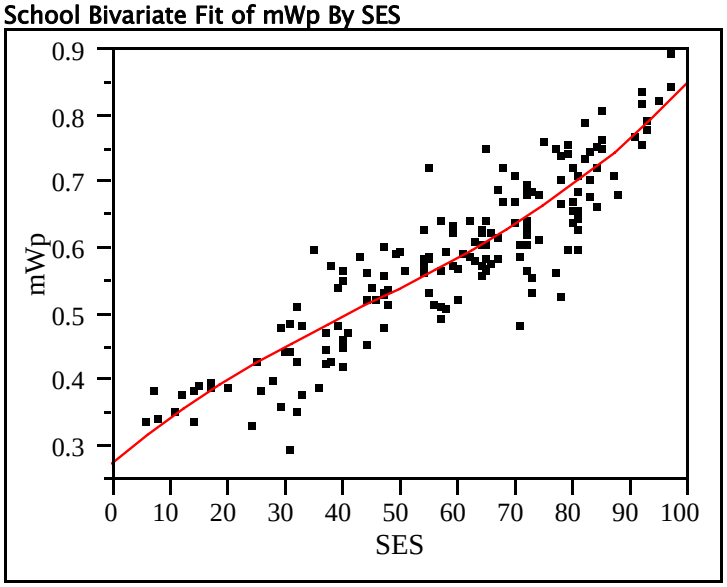
Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	3	1.1724582	0.390819	102.9450
Error	160	0.6074225	0.003796	**Prob > F**
C. Total	163	1.7798807		<.0001

(Continued on next page)...

Figure 2.

School Bivariate Fit of mWp By SES



Polynomial Fit Degree=3

mWp = 0.2989573 + 0.0047282 SES + 0.0000221 (SES-58.8393)^2 + 5.03e-7 (SES-58.8393)^3

Summary of Fit

RSquare	0.798756
RSquare Adj	0.795075
Root Mean Square Error	0.055864
Mean of Response	0.58506
Observations (or Sum Wgts)	168

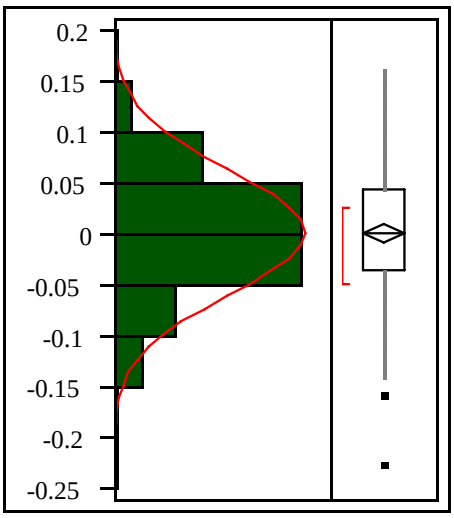
Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	3	2.0314481	0.677149	216.9777
Error	164	0.5118153	0.003121	**Prob > F**
C. Total	167	2.5432634		<.0001

(Continued on next page)...

Figure 3.

District Distribution



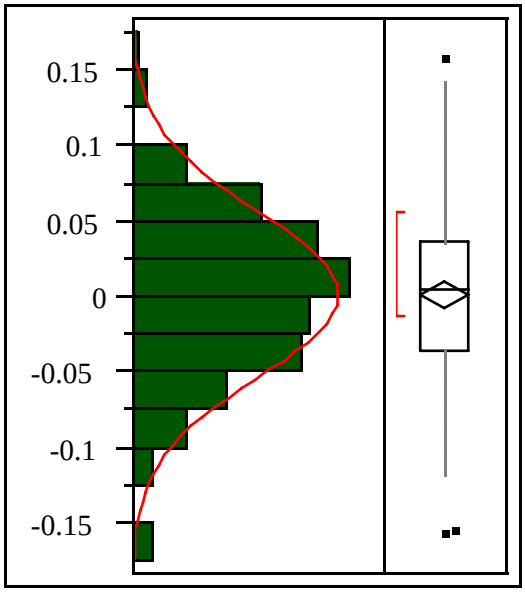
Normal

Moments

Mean	0.0002378
Std Dev	0.0610465
Std Err Mean	0.0047669
upper 95% Mean	0.0096507
lower 95% Mean	-0.009175
N	164

Figure 4.

School Distribution



Normal

Moments

Mean	0.000125
Std Dev	0.0553939
Std Err Mean	0.0042737
upper 95% Mean	0.0085625
lower 95% Mean	-0.008313
N	168

(Continued on next page)...

policy-makers a powerful tool for making comparisons between and among schools of discrepant SES. Districts and schools with scores of $6 < W < 12$ seem to be performing better than would be expected, while districts and schools scoring higher than 12 could be considered "exemplary".

Caution is still needed in interpreting these scores, however, as there can be other sources of error besides SES that can affect student outcomes on the WASL. For example, one school in the study scored at $W > 14$; however upon investigation it was found to be a magnet Talented & Gifted school in a large district. Another example of confounding error was in a small district with $W > 13$; this small district is home to a large university (many highly educated parents are often low-income graduate students).

Another cautionary note: Kurki et al. (2005) suggest that there may be better measures of poverty than percentage of students receiving free/reduced school lunch in terms of the strength of the negative correlation to school achievement. Unfortunately, most of these data are not easily obtainable.

The following school districts had mean scores above the second standard deviation for grades six, seven and eight in this study: Nooksack Valley, Cascade, Tonasket, and Pullman. The following schools within larger districts had mean scores above the second standard deviation: Washington Middle School (Seattle); Morris Ford Middle School (Franklin-Pierce).

FUTURE WORK

The middle grades were selected for this study because our district has been re-examining the effectiveness of our middle school programs, and we were interested to find possible exemplars that have perhaps served as laboratories for modern middle school research findings. We have been able to examine more closely the educational practices in exemplary middle schools and districts (those above the second standard deviation), which examination has already been helpful to our school improvement process. Using this method of calculating Educational Work, further study of schools and districts for grades 3-5 and high school should be fruitful.

REFERENCES

Kurki, A.; Boyle, A. & Aladjem, D.K. (2005). Beyond Free Lunch: Alternative Poverty Measures in Educational Research and Program Evaluation. *Am. Inst. Research. Pub.* 1-25.

Raudenbush, S. W. (2004). Schooling, Statistics, and Poverty: Can We Measure School Improvement? *William H. Angoff Memorial Lecture Report, 9.* 1-42.

Sirin, S. R. (2005). Socioeconomic Status and Academic Achievement: A Meta-Analytic Review or Research. *Rev. Ed. Research, 75:3.* 417-453.

White, K. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin, 91,* 461-481.

-Bruce Dean Kelley, Ph.D. is Director of School Improvement for Battle Grounds Public Schools

Stupid Excel Tricks for Assessment Folks

By Patrick Cummings

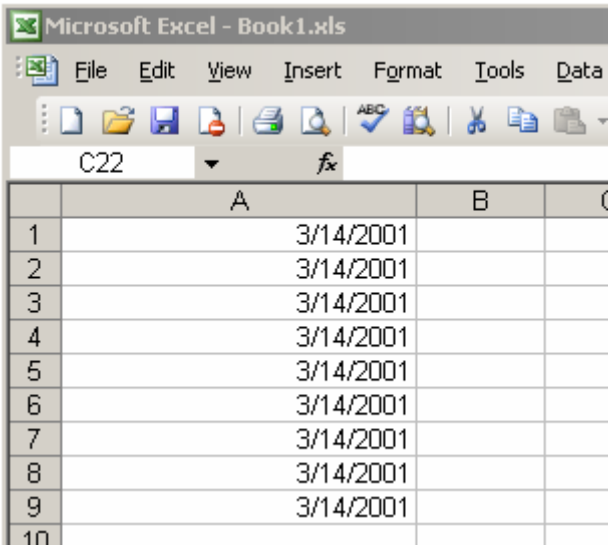
Stupid Excel Tricks for Assessment Folks

Introduction

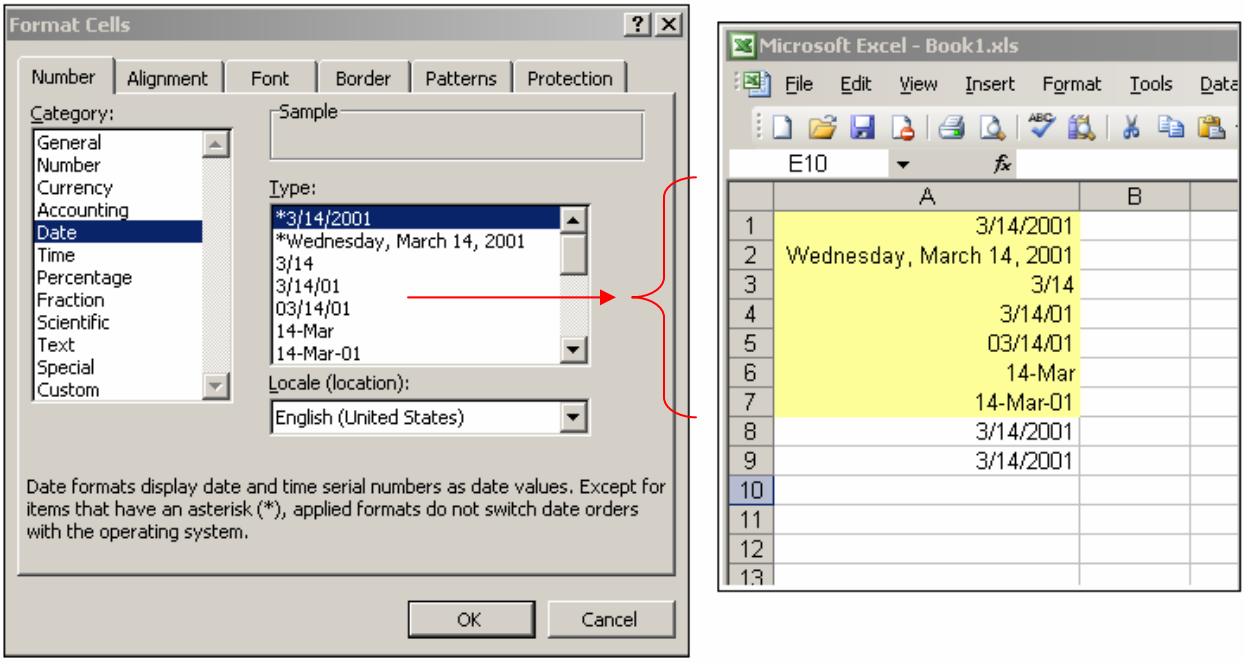
Have you ever stayed up late at night pondering how Excel handles **Date** functions? Me too. The following information should help you get a better handle on how Excel works with dates.

Formatting Dates

Let's start with a blank spread sheet. Enter the date **3/14/2001** (March 14th 2001) in the first 9 cells:



Just for practice, from the menu select **Format**, **Cell**, **Date** and change the first few dates to different format **Types**:

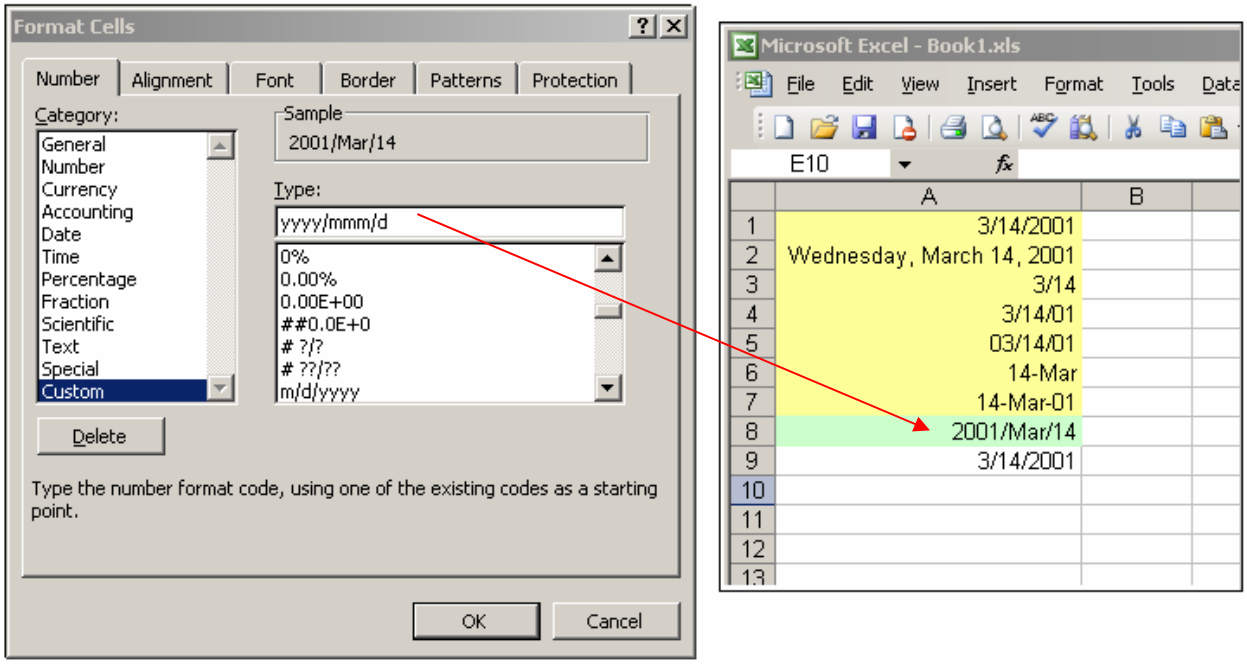


(Continued on next page)...

To continue with date formatting, let's assume that you need to customize the date in a way not listed in the menu. This can be done by selecting the **Custom** category. You can rearrange the year, month and day in any order you like. You need to know that year = y, month = m and day = d. Here are a few more details on formatting:

Year	yy = 01	yyyy = 2001		
Month	m = 3	m = 03	mmm = Mar	mmmm= March
Day	d = 14	dd = 14 lead with "0"	ddd = Wed	dddd = Wednesday

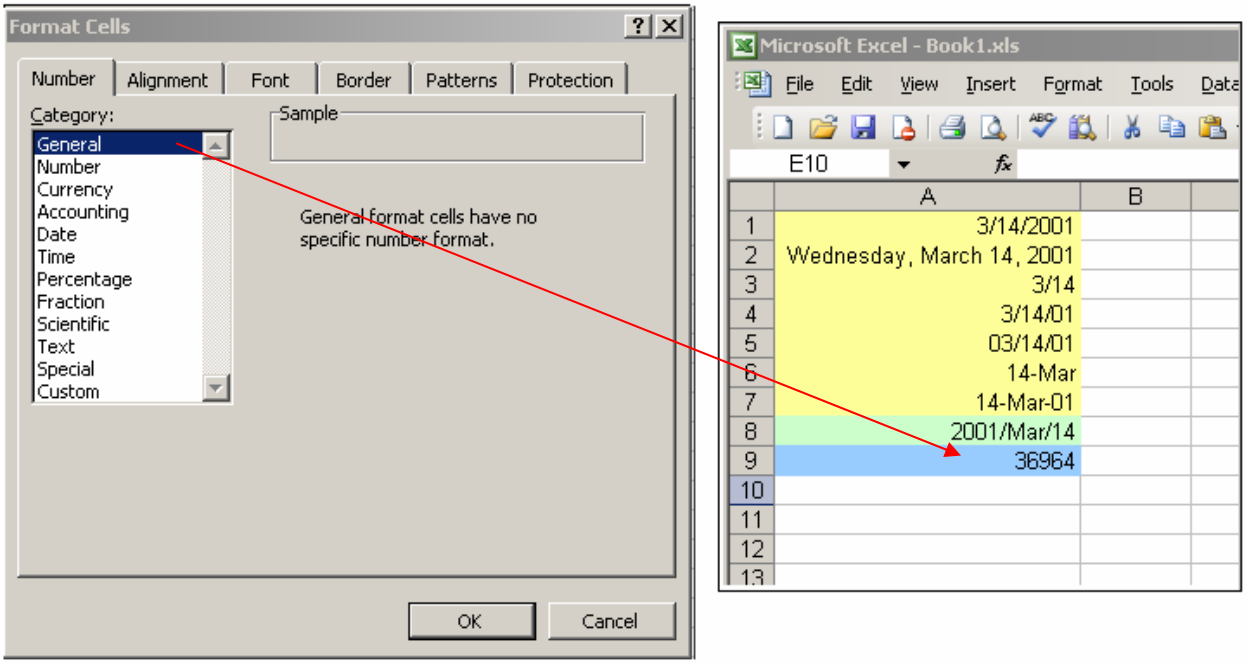
Here I have decided to customize the date in a year, month and day arrangement (**yyyy/mmm/d**):



(Continued on next page)...

Now the Date Secret

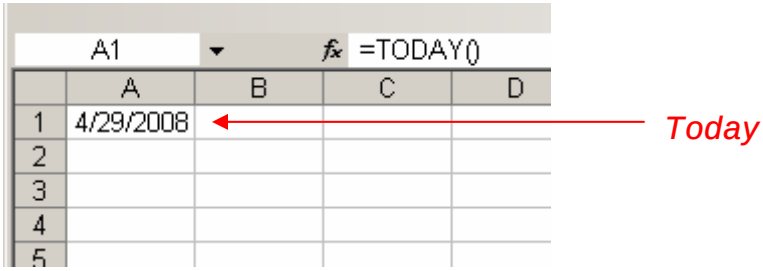
Up to now this has been relatively straight forward; enter a date and format in various ways either with a built-in **Date Type** or a **Custom Type**. The real secret is that Excel stores dates as numbers (specifically integers) or sequential serial number in the worksheet. The date **March 14, 2001** has a numerical value of **36964**. To confirm this, change the last date in your spreadsheet to a **General Category**:



To understand dates, you need to know that **January 1, 1900** has a value of **1**. By now, you should have figured out that the number presents the number of days the date is from January 1, 1900. That's right. **March 14, 2001** is **36,964** days form **January 1, 1900**.

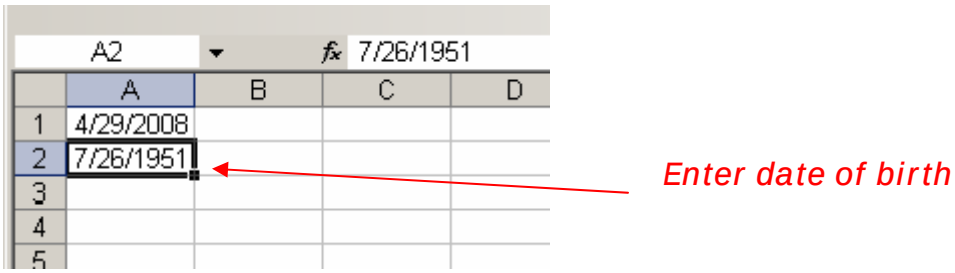
Next Steps

Now that you know Excel considers dates as numbers let's play a little. Start by opening a new spreadsheet. You can save yourself the effort of entering today's date manually. Excel has a function=**TODAY()**, that creates the current date. All you need to do is place this function in cell A1 and today's date is inserted:

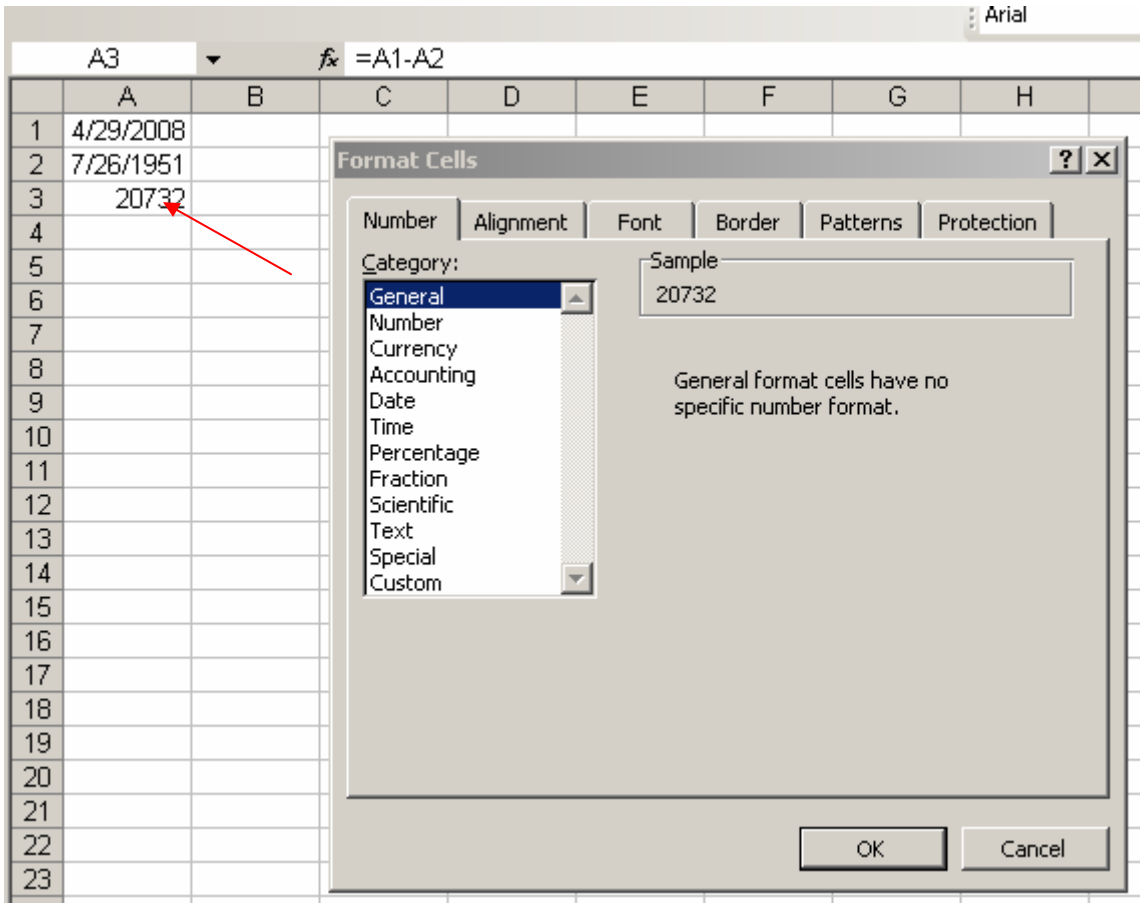


(Continued on next page)...

Next, in cell A2, enter your Date of Birth. I will use my birthday of July 26, 1951..... yes, I am a Leo.

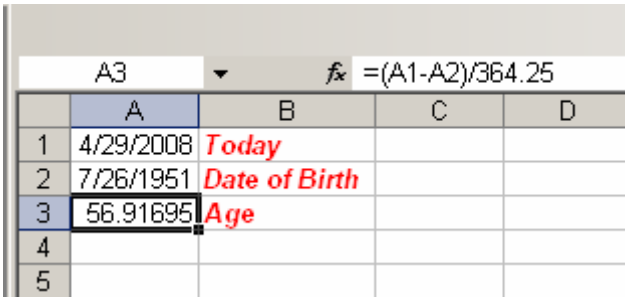


Now, in cell A3, subtract the **today's date** from **date of birth** (=A1-A2) and reformat to **General**. This means that I am **20,732** days old.

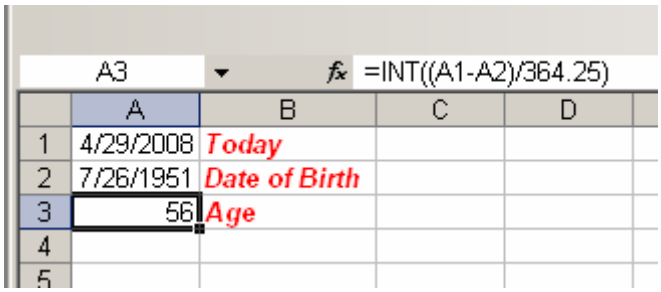


(Continued on next page)...

Next, we need to convert this number of days into a number of years. Most years have 365 days but every fourth year has 366 days. So the average number of years is 365.25. Let's modify our formula to **=(A1-A2)/365.25.** This makes me 56.9 years old.



To say that I am 56.9 years old is still not quite right. We are getting an accurate result but we don't really want to see the fraction. As a last refinement we'll wrap the whole thing inside an **INT()** function to give us a whole number (an **integer**). This is better than changing the number of decimal places displayed, which would risk some numbers being rounded up and giving an incorrect result. Here's the finished result...



Conclusion

Other than the fact that you now know I am 56 years old, you should also have learned that Excel considers dates as numbers. I hope you enjoyed another stupid Excel trick to add to your bag of tricks.

-Patrick Cummings is Director of Research and Evaluation for Tacoma Public Schools and is a regular contributor. Contact him at pcummin@tacoma.k12.wa.us

WERA Board Members

Nancy Arnold, President
Director, Special Programs
Puyallup School District
214 W. Main
Puyallup, WA 98371
253-435-6532
253-841-8655 (fax)
arnoldnl@puyallup.k12.wa.us

Emilie Hard, President-Elect
Principal, Glacier Park Elementary School
Tahoma School District
23700 SE 280th Street
Maple Valley WA 98038
425-432-7294
425-432-6795 (fax)
ehard@tahoma.wednet.edu

Lorna Spear, Past President
Executive Director of Teaching and Learning,
Spokane School District
200 North Bernard
Spokane WA 99201
509-354-7339
509-354-5965 (fax)
lornas@spokaneschools.org

Leonard Winchell, Executive Secretary
Washington Educational Research Association
PO Box 64489
University Place, WA 98464
253-564-4816
253-564-4816 (fax)
lenwwa@aol.com

Members-at-Large

Phil Dommes
Director of Assessment
North Thurston School District
305 College Street NE
Lacey WA 98516
360-412-4465
360-412-4555 (fax)
pdommes@nthurston.k12.wa.us
(Term expires April 30, 2011)

James Leffler
Program Director, Services to the Field
Northwest Regional Educational Laboratory
101 SW Main Ste. 500
Portland OR 97204
800-547-6339 ext 549
503-275-9584 (fax)
lefflerj@nwrel.org
(Term expires April 30, 2009)

Andrea Meld
Research/Assessment Analyst OSPI
PO Box 47200
Olympia WA 98504
360-725-6438
360-725-6333 (fax)
Andrea.meld@k12.wa.us
(Term expires April 30, 2009)

Paul Stern
Senior Research Associate
Washington State University -SESRC
PO Box 43170
Olympia WA 98504
360-576-6030
360-586-2279 (fax)
sternpo@wsu.edu
(Term expires April 30, 2011)

Washington Educational Research Association
PO Box 64489
University Place, WA 98464

We're on the Web!

Visit us at:

www.wera-web.org

The Standard Deviation
May 2008

The Standard Deviation is published spring, winter, and fall as an online newsletter. Submissions are welcomed from WERA members and others. Kindly submit articles for consideration using APA format