The

BEACON

News from

The Coalition for Excellence in Science and Math Education

Volume VII, No 2 **Copyright © 2003** **April**

In this issue: INTELLIGENT DESIGN, Bill MacPherson—CESE MEMBER JACK JEKOWSKI MAKING A DIFFERENCE, Jerry Shelton—STATISTICAL TESTS, Walt Murfin— SPECIAL EVENT on APRIL 13TH and TOON by Thomas

Intelligent Design

Periodic review of our New Mexico science standards (for K thru 12) has been under way for a while. The latest draft will be reviewed by the State Department of Education this summer. After this comes the point where the standards will be most vulnerable, when the State Board of Education will take public comment on the standards. That is, when the Intelligent Design Creationists mount their main attack on good science. In order to prepare our membership to deal with ID arguments, our sister organization, New Mexicans for Science and Reason (NMSR) has prepared a number of pages at their web site that address the main arguments of ID proponents. www.nmsr.org/iconanti.htm addresses the arguments of Jonathan Wells presented in his book *Icons of Evolution*. This web site has essays discussing many of Wells' positions on many of the famous "icons," such as Haeckel's embryos, the archaeopteryx, peppered moths, Darwin's finches, etc. This web page makes a wonderful resource for teachers attempting to counter arguments brought up in class by students who have read Wells' book, but are not very conversant with actual biology.

Another NMSR web page tackles the arguments of the "mathematical genius" of the ID movement, William Dembski. NMSR's arguments can be found at www.nmsr.org/dembski.htm

Everybody's favorite "irreducible complexity" guru, Michael Behe, is also addressed on the NMSR web page, and there are a number of other good sites that address Behe. Any search on Kenneth Miller will find a wealth of information on why there is no such thing as "irreducible complexity."

The NMSR web page has also addressed the "elder statesman" of ID, Phillip Johnson. One thing about Johnson is that he is very short on science and very long on argumentation.See www.nmsr.org/johnson.htm

In addition there is the Talk Origins archive, which has been around for years and is being constantly updated. They started out fighting "young earth creationism," and have kept up to date as creationism has evolved. The web page for Talk Origins is www.talkorigins.org This site was the first one I found when the creationism controversy came to New Mexico, and I still think it is the best.

Bill MacPherson
CESE President

The BEACON is published by the Coalition for Excellence in Science and Math Education (CESE). A 501(c) 3 nonprofit corporation, CESE is incorporated in the State of New Mexico. Visit the CESE web site.

WWW.CESAME-NM.ORG
David Johnson, Web Master

MEMBERSHIP INFORMATION:
CESE annual dues are $25 for individual, $35 for family, and $10 for students. Please make checks payable to CESE, and mail to 11617 Snowheights Blvd. NE, Albuquerque, NM 87112-3157

E-mail submissions (subject to edit) to **nshelton10@comcast.net.**

CESE MEMBERS MAKING A DIFFERENCE

Jack Jekowski

Jack's intense involvement with education reform began 30 years ago when he was working in Los Alamos for EG&G, and subsequently become more intense when he was given a "community outreach" assignment by his last employer, Allied Signal in 1996.

His work in the "military industrial complex" began while he was a student at Northeastern University, a college with a "coop" program. His work assignment was with EG&G, headquartered in Boston. That led to full time employment with EG&G and interesting jobs that involved travel to various parts of the country, and beyond.

Sometimes that travel included visits to the Land of Enchantment, and Jack was enchanted. (EG&G had offices in what is now the UNM/Industrial area near University Blvd. and Avenida Cesar Chavez, as well as at the Nevada Test Site, and opened offices in Los Alamos and on Kirtland Air Force base in the late 1960's. Along the way, corporate mergers resulted in EG&G's operation on Kirtland Air Force Base becoming part of Allied Signal, and later part of Honeywell Int'l.) Jack moved to New Mexico in 1970 and has never looked back.

From Jack's resume: "Mr. Jekowski was vice president and Sandia program manager for EG&G Energy Measurements, Inc. In that position he managed the 300-person Kirtland Operations in New Mexico and a 150-person operation in Las Vegas, Nevada supporting the former Defense Nuclear Agency and Sandia National Laboratories' nuclear weapons test activities, the Department of Energy's Transportation Safeguards program and other national security programs. Mr. Jekowski also managed engineering research and development activities for EG&G in Los Alamos New Mexico, supporting the national laboratory in scientific experiments world wide and at the Nevada Test Site." After retirement from Honeywell, Jack and two other EG&G alums founded Innovative Technology Partnerships, LLC in Albuquerque "to assist business, education, healthcare and government organizations to prepare for the uncertainties of the 21st century."

A day at the office

A recent day at the ITP offices included (federally contracted) consultation with representatives from Los Alamos National Lab. Their purpose was to lay out a plan to introduce a

"systems approach" to New Mexico's education establishment. This would include a feedback and evaluation phase. It seems hard to believe that, so far, such an approach is so lacking in our education system – across the nation. Visit their web site www.itpnm.com to get an idea of the scope of their activities.

One thing required to prepare for the uncertainties of the 21st century is improved education, particularly math and science education. Through an interesting combination of contracted consulting work and continuing community engagement, Jack has been deeply involved in analysis and assessment of our education system and studying avenues for improvement. He is active in more than twenty New Mexico community outreach and business organizations, most of which have a commitment to better education. Further, he is frequently a member of management in the organizations he is involved with. For example, upon passage of the national School-to-Work Opportunity Act (STW) in 1996, when the Albuquerque Business & Education Compact formed the Middle Rio Grande Business & Education Collaborative to implement STW, Jack became a Board member. He is currently the chair of MRGBEC.

When the Baldrige National Quality Program added education to its areas of interest, Jack became involved. He helped enlist the GBEEs (Governor's Business Executives for Education, started under Governor Bruce King's administration) as sponsors of an initiative called Quality Leadership in Education (QLE), to complement the existing successful Strengthening Quality in Schools (SQS), a Baldrige-based model to improve the use of quality principles in the classroom. Wherever SQS has been diligently engaged and supported over time (currently about 100 New Mexico schools), the improvement in student achievement has been little short of spectacular. For more about SQS, visit www.sandia.gov/sqs.

Baldrige expands

Enthusiastic support of the SQS initiative led to New Mexico's selection as one of six states to join a pilot program, Baldrige in Education (BiE IN), http://www.biein.org/, to imbed Baldrige principles in our education superstructure as well as in our individual schools through SQS. The State Department of Education, and many local school boards have enthusiastically participated in Baldrige training in the past few years.

Last summer, Jack accompanied a group of New Mexico SQS officials on a visit with gubernatorial candidate Bill Richardson to SQS school Georgia O'Keeffe Elementary (one of our New Mexico schools frequently visited by out-of-state dignitaries). One of the most persuasive testimonials came from a teacher at Tohatchi Middle School on the Navajo Reservation, where 58% of students are from low-income families, 70% come from homes where both English and Navajo are spoken, and 58% are Limited English Proficient. Once one of our state's poorest-performing schools, Tohatchi was named to our State Department of Education's (SDE) list of "Most Improved Schools" in 2000, and is now one of our best performing schools. The teacher reported that their success can be directly linked to their involvement with SQS. Richardson asked, "Why aren't all our schools taking advantage of this program?" Since then he has signed legislation restoring state funding to SQS that was twice vetoed by Governor Johnson. Jack estimates that an appropriation of about $3 million per year over a period of five years could bring SQS to every school in New Mexico. In addition, Jack is pleased that Richardson's administration sees the positive connection between an educated workforce and economic development.

CESE comes along

Jack's acquaintance with Marshall Berman began when they worked together on a Sandia Lab program (all the principal partners of ITP hold active "Q" clearances). He was at first reluctant to add yet another organization to his list of memberships, but eventually Marshall's persuasion won him over, and we are proud to have him as an active member of CESE.

Continued on page 4

Continued from page 3

Jack particularly enjoys the comfortable (electronic) give-and-take among the subducks. "Somebody can test an idea against a diversity of experience, knowledge and opinion." He enjoys reading, even when not participating. "It's a great forum for academic discussion."

Beyond that, Jack feels that CESE dramatically influenced the debate over education reform during the 2001 legislative session. This was primarily a result of Timothy Moy's presentation to a joint meeting of the House and Senate Education Committees. And, no surprise, Jack is a staunch advocate of data-based decision-making. He hopes our legislators continue learning to ask questions such as, "Where is your data?" and "What are the likely unintended consequences of the action you propose?"

How to "fix" education

After six years of up-close observation of our state government in action, much of what Jack feels is frustration. But he also sees reasons for optimism. Following are some of his ideas.

• Teachers need more help in the classroom, to allow time for such things as engaging individual students, training substitutes, consulting with peer groups, opportunities for middle school teachers to provide feedback to elementary teachers, etc. This would take money, of course, but experience in other venues shows that such strategies do work.

• There is great need for a more coordinated approach to education reform, rather than a series of "spot fixes" with little attention to unintended consequences in what is an extremely complicated system.

• There is much consternation over the supposed shortcomings of New Mexico schools. However, most of the suggested changes overlook the fact that the famous "700 pages" of rules governing our schools were created to implement the 350 tangled pages of statutes enacted by our legislature over a period of many years. Here lies fertile ground for improvement.

• Our system for handling the hundreds of proposed items of legislation places an impossible burden on legislators with respect to developing a reasonable understanding of what is being proposed. Even less are the possibilities for understanding ramifications and possible unintended consequences of proposals. There is great need of better pre-evaluation processes for proposed legislation before it is submitted.

• There is need for the many organizations interested in education to work together and develop a common strategy, keeping in mind that changing such a complex system takes more time than most people expect.

• A major underlying problem is that our educational system is still basically on the same track that existed in the early 20th century, while our society's needs have gone off on a different track. Serious re-routing is needed.

• Long-range planners in our federal government appear to be assuming a match between society's needs in the near future, and the talents and level of competence among graduates that probably won't be there. A significant long term commitment needs to be made by federal agencies to work with the K-16 system to ensure that an articulated and contextual learning system is created to ensure all of our students have an opportunity to participate in the 21st Century.

It seems reasonable to assume that Jack Jekowski will continue to be actively engaged in New Mexico, and national, education issues. We certainly hope so.

Jerry Shelton
CESE Treasurer

STATISTICAL TESTS

Does everyone remember the purpose of statistical tests? For every experiment we have an experimental hypothesis and a null hypothesis. Although we can't actually "prove" the experimental hypothesis, we can show that it is unlikely that the null hypothesis is true. Then, if the experimental and null hypotheses are exhaustive and mutually exclusive (meaning that's all there is, and if it's one it can't be the other), we assume that our experimental hypothesis might hold water, at least until some troublemaker comes along with contrary evidence. In the following, "p" is the limiting probability for rejecting the null hypothesis, conventionally 0.05. In other words, we say that the odds are 19 to 1 that the null hypothesis is false, and by inference, that the experimental hypothesis can be safely accepted.

Often the experimenter and the statistician are two different people. The experimenter knows a lot about the theoretical basis of the experiment. The statistician knows a lot about statistical tests. Often neither is conversant with the other's field and is not motivated to learn. This is an invitation to trouble. One solution is for the experimenter to learn enough about experimental design and statistical testing to be able to tell whether the statistician's advice is relevant. The very best situation is for the experimenter to be well enough informed to make the statistical decisions him/her self and just run them past a professional statistician for verification.

Lots of tests are available. The first rule of testing is to use the very minimum number of tests to get by. Each test has an inherent error rate. The more tests you run, the more data fishing you do, the greater the chance of an embarrassing error. We can't go through many of these tests in a short tutorial, so let's limit the discussion to a few tests. You understand that these tests have limited applicability, and you need to do some intensive studying to learn more.

Suppose our outcome is membership in a class – big vs. little – or less than 10, 10 to 15, greater than 15 – or black, gray, white; whatever. It is understood that we set up the categories in advance. We postulate the expected frequency or number of occurrences in each category, perhaps based on some physical law, and find observed frequencies to compare with our *a priori* expectations. We don't need to know a thing about the distribution of any variables. All we need is a table of the observed and expected frequencies. We can use the *Chi-squared Test* to determine whether our observed frequencies are reasonably close to our prior estimates. We calculate the Chi-squared statistic:

$$T(\nu) = \frac{(M_1 - M_2)}{\sqrt{S_1^2 + S_2^2}}$$

Chi-squared has a parameter (ν) called *degrees of freedom*: the number of independent ways our table of observed frequencies can be changed. If the number of entries in a row is c, and we know the sum of the entries, we could arbitrarily change only c-1 entries. The final "c-th" value would be completely determined by the total and the first c-1 values. If the table of observed values has r rows and c columns then ν = (r-1)(c-1). If either term in the parentheses would be zero, we strike it out. So if our table has 2 rows and 4 columns of observed values the degrees of freedom will be ν =(2-1)(4-1)=3. We can look up the value of χ^2 at whatever probability we choose for rejecting the null hypothesis. For example, at p= 0.05, $\chi^2(3)$=7.815.

If our experimental hypothesis is that the observed and experimental frequencies should be the same, we reject the null hypothesis if our value is less than 7.815. If the experimental hypothesis is that the frequencies would be different, we reject the null hypothesis if our value is greater than 7.815. Here's an example. You have a sample of 100, and you postulate that the sample is normally distributed with a mean of 50 and a standard deviation of 10. You set up a table, and count the number in each bin. This table has one row of observed values; it's a 1 x 4 table. Strike out the row term because (r-1)=0, so ν =(c-1)=3.

Continued of page 6

Statistical tests continued from page 5

Boundaries→	*<30*	*30 to 49.999*	*50 to 70*	*>70*
Observed→	*3*	*37*	*55*	*5*
Expected→	*2.3*	*47.7*	*47.7*	*2.3*

For this data Chi-squared (3)=6.90, which is less than 7.815, so we can say that the observed frequencies are close enough to the expected frequencies. Whoa! Isn't it obvious that the observed and expected frequencies are only vaguely similar? We got the results we wanted by the way we set up our bins. Use of chi-squared doesn't mean that we're lying, but it can make it easier. We ought to increase the number of bins and see what happens. The test means something if and only if we set it up honestly.

If we have two independent sets of data, and know (or are willing to assume) that both sets are close to normally distributed, we can use a *T-test*. Get the means of both sets, M1 and M2. If our experimental hypothesis is M1>M2, the null hypothesis is obviously M1<=M2 If the experimental hypothesis is M1<M2 the null hypothesis is obviously M1>=M2. Can you guess what null hypothesis corresponds to the experimental hypothesis that they are equal? The difference of the means follows a T-distribution. The value of the T-variable is:

$$T(\nu) = \frac{(M_1 - M_2)}{\sqrt{S_1^2 + S_2^2}}$$

The S's are the standard deviations, technically for the populations, but go ahead and use the sample values as long as the samples are nearly the same size. Values of the T-statistic are tabulated in many statistics books. If the sizes of the two samples are N1 and N2 the degrees of freedom are N1+N2–2. Given the known means, we could arbitrarily specify N1-1 out of the first set and N2-1 out of the second set. If we said only that the means are different, half the distribution could be in the upper tail and half in the lower tail, so we look up the value of the T-statistic for n degrees of freedom and p/2. If we said (in advance, of course) that M1>M2 or M1<M2 we can use the value of T at probability p. The calculated T is compared with the value of T from the tables. If we said "greater than", and our T is greater that the value from the tables at p, we reject the null hypothesis. If we said "different" we do the same, but use the value at p/2. If we said the means were equal, we reject the null hypothesis if our T is less than the value from the table. Think carefully: do you want to use "greater than" or "less than" as your test? Will the sign of T be important? You need to decide this for each case. Most textbooks don't explain this well enough.

This only works if we have just two sets of data. If we have a factorial experiment with several levels of each factor, we use the *F-test*. It's the ratio of two variances: an experimental variance and an error variance. The procedure is called Analysis of Variance, or ANOVA. If the experimental variance (the variance due to the imposed levels) is large compared to the error variance, we can reject the null hypothesis that there is no experimental effect. The actual calculation of F is beyond what we can cover in a short tutorial. There are restrictions on the F-test. Don't worry about them. The test is quite robust with respect to these restrictions.

The null hypothesis for the F-test is that there is no experimental effect. If we have two treatment factors, each with several levels, we will have null hypotheses of no experimental effect for either factor and for the interaction between factors. Then we'll need to test the differences between levels in each treatment factor. You can see how the tests can quickly get out of hand if you have several factors. Canned software is available for working out the math. There is a catch. You have to match the setup of the software to the design of your experiment; you have to know what you're doing. Unfortunately, every user is not as expert at this as one might hope. Sophisticated software doesn't mean it will be right!

This will not prepare you to carry out these calculations yourselves without additional study. However, when you run across these terms in a paper, you will at least recognize what they mean. You know what they are trying to do, and roughly how they do it. Maybe it will also help you to see if they are telling the truth. You might get the heretical idea that one purpose of statistical tests is to give dubious experiments a spurious cachet of respectability. We hope that's not often true but it does happen.

Walt Murfin
CESE Statistician

APRIL 13th, 2003 SPECIAL EVENT - JIM OBERG

*Space expert **Jim Oberg** will be in New Mexico the weekend of **April 13th.** He will give a public lecture on his analyses of the Apollo "hoax" mythologies. Oberg will also give us an update on the shuttle Columbia investigation. The lecture, entitled **"Lessons of the 'Moon Hoax Myth',"** will be held at the **LodeStar Planetarium at the New Mexico Museum of Natural History and Science, at 7:00 P.M. on Sunday, April 13th.** The lecture is co-sponsored by NMSR, by the New Mexico Academy of Science, and by the Coalition for Excellence in Science and Math Education. A small fee (still under negotiation, but definitely not more than $5 per person) will be charged.*

Jim Oberg

Oberg's web site: http://www.jamesoberg.com/

Study: Teachers should emphasize math concepts

From CNN.com/Education, Wednesday, March 26, 2003. See
http://www.cnn.com/2003/EDUCATION/03/26/math.countries.ap/index.html

WASHINGTON (AP) —American teachers must do more to help students understand the concepts of math, not just the mechanics of how to solve an algebra or geometry problem, an international review of 8th-grade classes suggests.

The four-year study analyzed videos of teaching techniques in seven countries, including six that score higher than the United States in math achievement: Japan, Hong Kong, the Czech Republic, Switzerland, the Netherlands and Australia.

Researchers cautioned Wednesday they cannot draw direct links between the teaching techniques and the countries' levels of success, particularly because so many factors affect learning. They also said the study did not aim to pinpoint features of good teaching, although more view is planned that could help produce specific tips for U.S. teachers.

The authors said U.S. teachers spend less time than counterparts in higher-achieving countries on explaining math's underpinnings.

"They're more focused on getting the answers, and less focused on giving students the opportunities to really engage in serious mathematical work," said James Stigler, chief executive officer of LessonLab, which conducted the study for the Department of Education.

"Finding ways to engage students in conceptual thinking—it doesn't fit within our cultural script of how you teach a math class," he said.

How those specific skills are developed "may be the real key," said University of Delaware professor James Hiebert, another leader of the math video study. Even when they present problems that link ideas to formulas, U.S. teachers often end up in step-by-step mode.

"We have to worry about whether students are understanding what they're being asked to do," he said. For example: "Why is that skill working? Why do you divide now? Why do you take the square root here? Why am I finding the lengths of the these two diagonals? The study underscores there is no single correct way to teach math, officials said.

It shows there are many paths to excellence in teaching," said William Frascella, an education leader of the National Science Foundation, a partner in the study. "Unfortunately, it appears from initial results that the United States is not able to use any one of these paths in a consistent and sustained manner."

The Education Department plans to make available public copies of videotape examples in compact disc form. Results of the science-video portion of the Trends in International Mathematics and Science Study will be released later.

The Coalition for Excellence in Science and Math Education
11617 Snowheights Blvd NE
Albuquerque NM 87112-3157

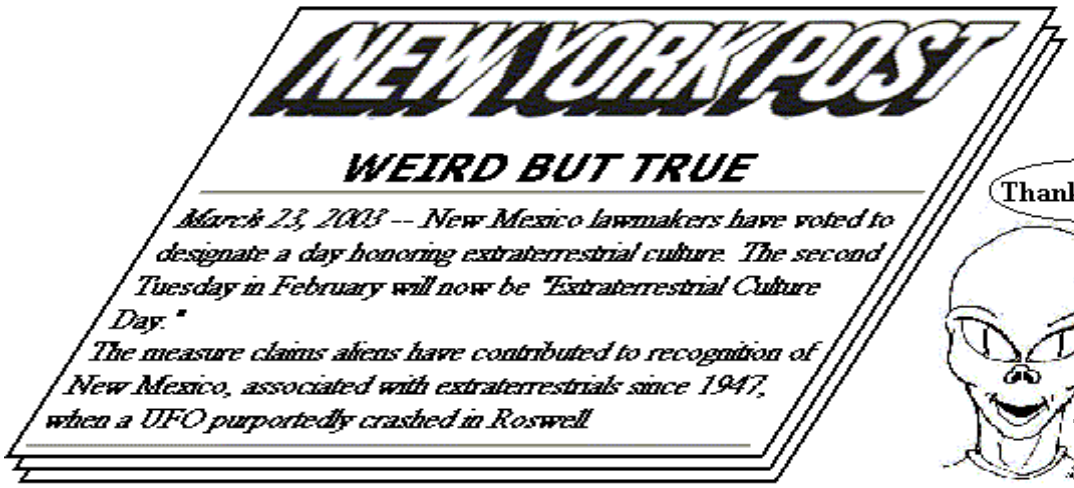
RETURN SERVICE REQUESTED

E.T., Phone Home - You've Got a Holiday!!

Dan Foley

On March 21, 2003, House Memorial 44, which names every second Tuesday in February as 'Extraterrestrial Culture Day,' passed 60-0 in the New Mexico Legislature. Sponsor Rep. Dan Foley, R-Roswell said the measure "helps advertise New Mexico."

And he's right! The measure was mentioned in dozens of news outlets in America and around the world, including "Weird News" from the Anchorage Daily News, & "Weird but True" in the New York Post. The Montreal Gazette of Canada described it as exemplary of the "Majesty of the law," and recognized the NM House as "that august body."



Toon by Thomas