



Women in Science and Engineering: What the Research *Really* Says

A panel discussion co-sponsored by WISELI and the Science Alliance.

Thursday April 14, 2005

Biotechnology Center Auditorium, University of Wisconsin – Madison

Panel discussion with Paul Whalen, Linda Oakley and Caitilyn Allen

C.A.: So the question was, do we know that the ACT is a good predictor of success in college, a better one than the SAT? Because the SAT has just been rewritten to more closely resemble the ACT. But I don't know.

Audience: This is a question I've been asking for many years and I haven't received a satisfactory answer, maybe Caitilyn you can help with this, what is the profile of the people who write the math tests?

C.A.: Actually it's not as trivial of a question as it might seem because they keep it a secret. And the Educational Testing Service is an interesting outfit to operate because they put all [of] their test-score data out, they're very open about the outcome, but the input and the research that goes into it is a very jealously guarded secret. And I did contact them and ask them about the gender distribution of their test-question writers in the math and they said that they don't discuss that.

John Wiley (audience): I'm going to say a little bit more about this later, but in answer to the previous question about the SAT or ACT as a good predictor, it's about as good a quantitative predictor as we have. But if you combine ACT or SAT and high school grades and class rank in the most defensible way you possibly can, I won't go into what that means, but you combine all that information in the best way possible it still explains only about twenty percent of the observed variance in first year college grades, which means eighty percent is due to something else. So the very best predictors we have all put together are pretty lousy. Secondly, I used to in fact work for ETS. I was on the board of directors for the engineering exam and so I've seen that process of putting tests together up close and I was always struck by the fact that there really was efforts made to make sure each group gets a question. It's rejected if it shows any gender or ethnic difference. They pre-test all those questions extensively. And any question that proves a systematic difference by gender or by race is rejected, never gets on the test at all. But when you put them all together and make up a full test, the differences pop out again. So it isn't in the individual questions, it's something in the experience of sitting down and mindlessly taking a number two pencil and filling those in. Some people find that interesting because it ...

Audience: Okay, so in your talk Linda you talked about the role of exercise or voluntary exercise for relieving stress. So would you predict then that women in my generation that grew up with Title XI will maybe have a different experience than women who grew up before that? ...

L.O.: The question was if voluntary exercise in the social competition situation is there a generational effect perhaps? Well, interestingly I just focused on [some]one who was studying this phenomenon and trying to get exercise knowing that it's a very inexpensive cure-all for many social illnesses. They did look at generation effects with women and the women that were picking it up and doing it without spending money and doing it in social groups – and it wasn't based on income these were blue-collar women as well as academic women as well as other[s] – and you saw more of a generational effect older than you saw a generational effect younger. But yet the messages were different. So I don't know. It's a good question.

L.O.: The question was to clarify a little bit more about the disparities, the sex disparities in depression, which is a very well characterized stress-related illness. No matter how you calculate it: by age, income, education, there's a two-fold female to male increase in depression. And it doesn't matter how you measure it, that holds.

Audience: But is it based on people reporting it themselves?

L.O.: It doesn't matter if it's objectively assessed, if it's self-reported, whatever parameters you use. They've double checked that statistic many, many times over. Because it's been around for a long time, everybody's been trying to write that equation to show where it doesn't work. But it's consistent. I think what's more interesting is that, at least if you get away from the gender differences in illness and you look at disparities,

then you're able to see that if you create the same environment – for example when they compare the heavier aggressive animals with the cooperative lighter-weight animals and put them both in the same defeat – everything else is the same. So that again it may well be that that biology, the endocrinology, doesn't really hold that much. Now the one paper that I did measure they did – because they were curious about that – they did measure testosterone responses as well as aural cortisol responses but even the testosterone responses in competition situations didn't tell you anything. So it may well be that as our science gets better that two-fold might go away, but right now the science always shows that consistently women are that disparity.

P.W.: My entire field at the moment, the fMRI was invented in 1992, any question you ask is about seven pages deep and with all the very small Ns, I think what's amazing is that how remarkable it's all working out. We're really finding everything we thought we'd find. And I think that should be the really upsetting part and the thing that tells you that you know you're seeing what you believed you were going to see because that's the horse-blinder you have on and that's what you're expecting. And there's probably, you know, there's a paradigm shift coming. I won't be alive for it, that's okay. That's what it's going to take because right now it's just working out really well. So, one point I wanted to make was if you believe women are more emotional, if you head into a brain analysis where you're taking the entire volume of the brain, slicing it up into 700,000 boxes, each one of those ends up being a t-test. We call that experiment-wise error and what that says is that by random chance you were going to find six cool stories so one of them could be the one you went in looking for and then you make up the other five.

Audience: What if you combine say the functional MRI study with Positron Emission Tomography and a whole series of other types of imaging methods ...

P.W.: I think combining them with other measurements, I don't know about other brain imaging techniques, but certainly there's a hundred and fifty years of psychological empiricism that is getting ignored as we reinvent the wheel and get *Nature* papers out of claiming that something's new when it was shown fifty years ago but just not with brain imaging. So, you're right if we take – these studies do not exist in a vacuum, if you can take them and tie them to the animal studies, the wealth of data that sits there, if you can tie them to some of those ninety-seven-cent paper-and-pencil tests, then you don't need a 600-dollar brain scan. You can come up with a neuro-signature. That's kind of like being a dentist, if I tell people this we'll put ourselves out of business. But really, that's what we do. We put people with anxiety in and we give them paper-and-pencil tests. The idea is, we're trying to predict who's going to do well on drugs, predict who at the age of twelve is going to get depression [and who] at the age of thirty-six. And if we can just use the right paper-and-pencil test and the right psycho-phys. to correlate with the brain data then in ten years we won't need to give everybody a brain scan to know about an effective treatment. We can use the paper-and-pencil test that is correlated with the brain scan ... So that's the goal, to put ourselves out of business by coming up with excellent correlates that don't cost as much.

P.W.: Can I ask you a quick question?

C.A.: Sure.

P.W.: I really liked the Wainer and Steinberg thing, the bidirectional validity thing. So, it's cool that you see performance in a college class and you see that the test didn't predict it well because they're thirty-three points off. But that thirty-three points is strikingly similar to the disparity in the mean itself.

C.A.: I noticed that too.

P.W.: So wouldn't you say that if you adjusted the math scores for performance in a college class, which probably accidentally controls very nicely for income and IQ and other things, you could really say a female 600 is a male 635? You haven't really been talking about anything all along, right?

C.A.: You could say that but –

P.W.: If performance in a college math class is the goal.

C.A.: Performance in a college math class is the ostensible goal but you'll notice that Lawrence Summers for example is using achievement on the high school, the test you take in your high school senior year to predict whether or not you're going to win the Nobel prize when you're fifty-two or bringing a bunch of nice big NIH grants to your

institution once they hire you as an assistant professor. So it wasn't really designed to do that, in fairness. But I also think that the psychological load, if you will, of having women know that they're going into the test and they're going to get thirty-five points on the way out the door just for being cute is no gift. In the long term it's not going to be a good strategy even if it's empirically respectable.

P.W.: No, no, no. I agree. You would have to fix the test based on past performance but it wouldn't be a raw scale, you wouldn't do it after the fact. The thing would be adjusted so that person would end up with a 630.

C.A.: [That would go] straight to court, straight to court!

P.W.: Alright.

L.O.: Thank you. The question is about enriching environment to try to counter the immobilization effects from social defeat as a way of competing when you're in a disadvantage. And these were animal studies and so it doesn't take much to take an animal from an essentially wire, food and water environment to an environment that sort of conveys positive things about the animal. So that there's something warm, there's something fuzzy, there's something – probably easier to explain the opposite environment, would be more of an institutional environment. Something like maybe most of our offices.

C.A.: I'm thinking about my own lab. Do I need to install carpeting?

L.O.: Well, I think if you look at the – to just push ahead quickly on your comment, an environment that discounts everything about you would be the opposite of an enriching environment. So if [you] come [as a] female and everything is male, if you come in black and everything is white, you come in speaking Spanish and everything is English, just that environment alone will have a way of defeating you, the individual, socially not intellectually, or not ... these exams. But we're talking about a competition; we're not just talking about how smart I am. Can I compete with my smarts here? And I think that that's what makes it very interesting to me. Thank you.