Good afternoon Rector Treacy, President Sands, members of the Board, Provost Clarke, and all others gathered here today.

We live in an age of data, where it seems every aspect of our lives is subject to quantitative measurement and analysis. Higher education is not exempted from this, and the life of faculty is more and more defined by obtaining, analyzing, and being evaluated by numerical metrics purporting to provide an objective numerical measure of nearly every aspect of our work.

Now, numbers are just carriers of information. They reflect whatever flaws and biases are present in the human-devised processes that produced them. But for some reason, numbers have a seductive quality of detached objectivity and authoritative precision that makes it easy to avoid asking hard questions about what they actually measure and whether they have sufficient reliability and validity to be used for any purpose at all. This pro-numerical bias is quite pervasive, even at a university full of numerically savvy people who know how to do these things right.

Let me give you just a couple of examples.

At the end of each semester, students are asked to fill out what is a known as a SPOT survey (for Student Perception of Teaching) where they evaluate each of their instructors. The SPOT survey has many components, but the most important is the question that asks students to rate their instructor’s overall effectiveness. This score is frequently the main vehicle by which faculty teaching is judged in annual evaluations, and therefore has a real impact on our salaries.

The problem is that what SPOT scores actually measure is student satisfaction. While this certainly has value, it is not the same thing as teaching effectiveness. It then comes as no surprise that some of the things that figure into student evaluation scores are how well dressed instructors are, how good jokes they tell, and, most importantly, what grade students expect to receive. Female faculty and faculty who speak with an accent consistently receive lower scores than their colleagues, and one of the best ways to lower your student evaluations is to make your course more challenging and ask students to work harder. These distortions are well known, yet SPOT scores continue to play a primary role in faculty evaluation simply because of convenience and probably also the seductive quality of a numerical score.

Another metrics-heavy area is student learning assessment, which is the name we give to methods designed to measure how much knowledge and skill students have acquired during a course. Accreditation agencies are increasingly demanding that universities use learning assessments to improve their instructional program. The problem with this is that there are very few documented examples of learning assessments actually leading to improved instruction. As outlined in this article (David Eubanks, “A Guide for the Perplexed,” Intersection Fall 2018, pp 4-13), doing assessment right requires specialized (and therefore scarce) expertise, more time
and resources than is available to the average faculty member, and sample sizes that are larger than most courses provide. Short of this, learning assessments are prone to be inconclusive or downright misleading. The rigor required to avoid this is virtually never applied to learning assessments, and yet we are told that they are crucial, and we need to do more of them.

I can come up with several other examples in this vein, but since my time is very limited, I thought I would balance my examples by pointing out aspects of University work where more metrics would in fact be beneficial. As an example, I would like to spotlight the recently released Electronic Faculty Reporting System, or EFARS, although I would like preface my remarks here by stating that I’m not trying to single anyone out, it’s just that EFARS offers a recent and instructive example of a much broader issue.

As faculty, we are required to report annually on our professional activities in areas such as teaching, research, and service. Until recently, this reporting was entirely paper-based, and the purpose of EFARS is to bring it into the electronic domain, partly so that the faculty’s scholarly and other work can be more easily quantified—and turned into metrics. Unfortunately, EFARS was rolled out before it was ready for primetime, using categorizations that were inappropriate for many faculty, and was weighed down with a very cumbersome interface. Many of us experienced a doubling or tripling of the amount of time required for activity reporting, which across the university has added up to many thousands of hours of wasted faculty time, corresponding to hundreds of thousands if not millions of dollars in lost productivity.

EFARS is now in the process of being fixed, but if Virginia Tech had a mechanism in place for measuring the time cost for the routine administrative procedures such as faculty activity reporting, this problem would have been caught before it caused any damage. Somewhat ironically, such data would be much simpler to obtain than reliable measures of student learning or the impact of scholarly output.

That doesn’t mean that we should give up on measuring such things. As a scientist, I believe in data. I believe it is possible to ask well-posed questions about how reality is put together and use the answers to enhance our understanding of the world and by extension improving the way we lead our lives. But I also know that doing this well is hard, and requires much more of us than simply going with the flow.

Given the pressures of today’s environment, it is clear that Virginia Tech is all but destined to become ever more data-driven with the passage of time. And as the university undergoes this transition, I believe it is vital that we commit to letting decisions be guided by high quality data that is well understood and used only for the purposes it is designed for. If we don’t do this hard work, we risk at best wasting resources on meaningless data collection tasks, and at worst creating perverse incentives that push the university in directions that are detrimental to its core mission. In other words, we need to do this right, or we need to not do it at all.

The good news is, as I just mentioned, that we have a university full of savvy people who know how to do this sort of thing right. Metrics and benchmarking is an issue that matters to faculty, and where we have all the requisite competency to make a difference. So it seems to me that this is an ideal subject for shared governance, and my purpose here today is to ask for a partnership with the administration to help answer questions on, of instance, the proper design and use of
SPOT surveys, how to best use data to improve instruction, and how to measure scholarly impact without discouraging risk taking and creativity. If you take us up on this offer, I think you will find faculty more than eager to assist you in finding our way forward in an ever more data-driven and metrics-reliant world.

Thank you.