American Physical Society "April Meeting"
Scientific Integrity in Government
Tampa, Florida
April 17, 2005

John Marburger
Director, Office of Science and Technology Policy
Executive Office of the President

Thanks to Lawrence Krauss for twisting my arm to appear on today's panel. He has done a fine job of bringing the beauty and excitement of physics to a broad audience. Perhaps the steady increase of high school students who take a physics course owes something to his effort. Greater awareness of quantitative science can only help in broader discussions of the interaction of science with society.

Science is not possible without very high standards of integrity no matter who does it or where it is done. I am not aware of any disagreement on this point, but as a scientist in government I am pleased to speak this afternoon on the issue of integrity.

Perhaps because it is so respected, science – or at least the *word* "science" – is appropriated by many groups in society for many purposes. Sometimes it is used to signal something dark and perhaps out-of-control, as in Judge Richard Posner's recent book "*Catastrophe*" where he refers to "maverick scientists" determined to put the world at risk with their particle accelerators. More often science is recruited for positive purposes in support of a position or a product – a dietary supplement, say, or an assertion about the consequences of chemicals in the environment. I think we should begin by being clear what science really is, and what aspects of it need integrity. For an audience of scientists, I can be brief on this topic.

I take science to refer to two kinds of activities, first, the development of *explanatory frameworks* through observations of nature, and second the development of *effective strategies* through empirical tests. Science products of the first kind include plate tectonics, quantum mechanics, and the Standard Model. The second kind includes drugs and therapies, electric motors, and integrated circuits. We use words like basic science for one and applied science for the other. The old conception that these lie along a one-dimensional spectrum that starts with basic research and ends with widgets is being replaced – not rapidly enough – with more realistic pictures that view the two aspects as separate dimensions in a multidimensional framework. This is what Donald Stokes' increasingly popular book "*Pasteur's Quadrant*" is all about.

However you slice these things that science does, what they all have in common is *empiricism*. That means either we ourselves, or others on our behalf, actually look at reality and not just speculate about it. If we do not do it, then we have to rely on others to observe and report the actual behavior of phenomena in a manner as independent as possible of any strategy or explanatory framework being tested. The quality of "actuality" is essential, and this is the first place where integrity comes in – I call it the *truth dimension* of integrity. We have to trust the reports of others, or repeat their work. Scientific reporting has evolved traditions to increase our confidence in these accounts. The precise conditions of observation need to be described in enough detail to permit readers, at least hypothetically, to make the observations themselves.

This kind of detailed reporting also allows us to judge the *diligence* of reporters and to conclude, for example, that the methods described are likely to capture all the relevant variables and track down all the relevant data. Diligence is so important to the usefulness of science that it deserves to be regarded as a second dimension of integrity. (My school teachers would agree.) When I was actively reviewing papers for technical journals, a common reason for a negative review was the author's neglect of others' work.

I mentioned reporting data "in a manner as independent as possible of any strategy or explanatory framework." This is a chicken-and-egg issue. We need a hypothesis to organize data collection, but the data are needed to support the hypothesis. In the resulting tension, the testimony of nature takes precedence, and this leads to a third dimension of integrity in science. We expect scientists to modify their observational strategies in response to what they find, and not vice versa. The worst violation of this principle is ignoring data that do not fit; more subtle is a stubborn resistance to completing the cycle of the scientific method which is to abandon or change the hypothesis in the face of evidence. We need to be aware of our possibly hidden assumptions in the first place, and be prepared to admit we guessed wrong. I view the mental attitude appropriate for this behavior as a third dimension of integrity, which might as well be called the dimension of *humility*. Attitudes toward humility and its opposite, arrogance, differentiate sharply between scientists and non-scientists. Scientists feel humble before nature, but express confidence in hypotheses tested by the scientific process. Non-scientists perceive that confidence as arrogance, which is not completely unreasonable given the provisional nature of all hypotheses.

To sum up, at least three dimensions – *truth, diligence,* and *humility* – are required to measure the integrity necessary for science.

The empirical aspect of science that makes integrity necessary also entails a frustrating quality of incompleteness. Physicists are lucky to have a well-defined field with the most stable theoretical structure in science. Most other fields are not like this. Controlled experiments in social science, for example, are possible but difficult to relate to "real life" situations in the same way as experiments in physics. Real physical systems like rockets and electrical machinery often permit a separation between a relatively few degrees of freedom that can be treated in detail from all the rest of the environment which is treated as a 'reservoir' with just a few parameters like temperature and pressure. Human situations have an enormous number of uncontrolled but significant variables, and extracting significance from human behavior in real life is notoriously difficult. Consequently there are many overlapping theoretical structures in the social sciences which do not enjoy the same degree of consensus as in physics, and the power of predictions is much weaker.

The predictive successes of the physical sciences contribute to a widespread misconception about the power of science to answer "what if" questions. This is important to our subject, because when the word "science" is recruited to a scientifically uncertain proposition, doubters are at a rhetorical disadvantage. The integrity inherent in the word "science" implicitly denies the quality of integrity to those who doubt the proposition. This makes the successful recruitment of the word "science" to a cause a powerful device for advocacy. Advocacy is not science because it ignores the provisional nature of the hypotheses it espouses. This is Michael Crichton's argument against the concept of consensus in science that he expounded in his Michelin Lecture at Caltech two years ago. I agree with Crichton that " In science consensus is irrelevant. What is relevant is reproducible results."

Consensus by itself does not improve the quality of prediction. In this connection, I would like to draw attention to the book "Prediction: Decision-Making and the Future of Nature" that evolved from a project in the late 1990's supported by the National Science Foundation and sponsored by several organizations including the Center for Science, Policy, and Outcomes which was then at Columbia and is now at Arizona State University. The project description reads as follows:

> "Prediction in traditional, reductionist natural science serves the role of validating hypotheses about invariant natural phenomena. In recent years, a new type of prediction has arisen in science, motivated in part by the needs of policy makers and the availability of new technologies. This new predictive science seeks to foretell the behavior of complex environmental phenomena such as climate change, earthquakes, and extreme weather events. Significant intellectual and financial resources are now devoted to such efforts, in the expectation that predictions will guide policy making. These expectations, however, derive in part from confusion about the different roles of prediction in science and society. Policy makers lack a framework for assessing when and if prediction can help achieve policy goals. This project is a first step toward developing such a framework."

This book should be better known because it exposes a misunderstanding about the role of science in topics that are controversial, and about which issues of integrity are sometimes raised. Predicting for hypothesis testing is profoundly different from forecasting for policy. The first kind of predicting is necessary to science, the second kind is not. The first kind applies to at least partially controlled experiments, the second must be applied to uncontrollable reality. The first is rarely controversial, the second is rarely not controversial.

Before I proceed, I had better make it clear that as a scientist/administrator I attempt to apply the values and lessons of science to everything I do. From this perspective the significance of all human actions is to be found in provisional hypotheses whose validity we pursue with the same qualities of truth, diligence, and humility that we apply to the conduct of science. I think there is a science – a social science – of science policy, and to practice it one must gather and report data objectively, labor diligently to discover all relevant factors, and be prepared to consider alternative interpretations of what we observe. The skepticism I believe is necessary when considering the significance of policy-relevant data leads me to question many assertions that I might otherwise accept if my role were advocacy and not counsel.

Having laid out my thoughts on science and integrity, let me turn to the roles of government in science. I can identify five: government is a *sponsor, performer, consumer*, and *regulator* of science and a *disseminator* of scientific information. I will comment briefly on aspects of integrity in each role.

The integrity of information *disseminated* by the federal government is subject to the Data Quality Act (or Information Quality Act) of 2000. This is a relatively recent law that should be better known. It "requires federal agencies to issue information quality guidelines ensuring the quality, utility, objectivity and integrity of information that they disseminate and provide mechanisms for affected persons to correct such information." "Integrity" in this context refers to the security of the data. "Congress' intent," according to the Thomas website, "was to prevent the harm that can occur when government websites, which are easily and often accessed by the public, disseminate inaccurate information."

In 2003 the President and I were sued under the Information Quality Act by an advocacy organization to prevent OSTP from disseminating the so-called National Assessment report on the consequences of climate change. We settled with the organization by appending a statement to the report on the OSTP website asserting that the document had not been subject to the data quality provisions of the Act. Someone unfamiliar with the history of this disclaimer may wonder why it is there and include it as evidence for a hypothesis about Administration attitudes toward climate change. This would be an example of a hidden assumption that undermines the face value of the evidence.

The Information Quality Act has been interpreted as also applying to the government's *regulatory role*. The Office of Management and Budget (OMB) has tried to improve the quality of data used in regulation by insisting on standards for agency peer review. Their initial attempts to do so met with criticism from the science community because the proposed guidelines were perceived as being too broad, covering traditional peer review procedures well beyond the rule-making process. The concerns were worked out in the rule-making process for the OMB proposal.

Agencies make rules in response to laws whose language is often the result of compromises needed to get the bills passed. Different parties to the compromise can have different understandings of what the words mean, and this leads to controversy during rule-making. Laws like the Endangered Species Act and the Clean Air Act are not unambiguous, and can be interpreted differently by different regulatory personnel. These acts operate in domains of phenomena where the science is difficult and suffers from the uncertainties I described earlier. They also deal with the emotionally charged issues of environmental quality and public health, and they have very strong economic impacts. Actions and events associated with the implementation of these and similar laws occur in a heated advocacy environment, and should be interpreted with great caution. The probability of incorrectly using words like "science" and "integrity" here is very high.

As a *consumer* of science, the federal government relies on advisory panels whose structure and operation are subject to the OMB peer review bulletin mentioned above, and by the Federal Advisory Committee Act of 1976 (FACA). The OMB bulletin applies to "influential scientific information" and "highly influential scientific assessments," the influence here being on governmental decision making. For these categories, agencies are required to engage in peer review best practices in selecting and using panels. As an alternative, agencies may use the National Academies to provide advice or reviews, which is a strong testimonial to the high degree of confidence the government has in the integrity of the National Academies processes.

The FACA provisions place very strong constraints on the appointment and operation of advisory panels, particularly on the appointment of members and the openness of the proceedings. These provisions apply to essentially any advisory group that has at least one non-government member. Language in the FACA make it clear that one of its purposes is to ensure the integrity of the advisory process. For example, it requires any legislation setting up an advisory committee to "contain appropriate provisions to assure that the advice and recommendations of the advisory committee will not be inappropriately influenced by the appointing authority or by any special interest, but will instead be the result of the advisory committee's independent judgment;" This is similar to scientific integrity because it seeks to preserve the actual content of the external advice. Whatever the quality of the advice, FACA says the officials receiving it, and the public they serve, need to hear it without adornment.

As a *sponsor* of science, government must judge what to support and be accountable to the public for expenditures. Priorities for support are driven entirely by societal values, expressed through the machinery of government. Most federally sponsored research addresses some societal need such as health, security, or energy. The most important parts of science policy formation are to determine what research is likely to have the greatest impact on these identified needs, and how much public money should be spent on it. This is a problem of cost/benefit analysis, which can be difficult to do for many types of science. To the extent that cost/benefit analyses are science, the three dimensions of scientific integrity apply. The assessment of benefits, however, involves many non-empirical judgments that are subject to political, not scientific, determination. Disentangling the roles of science and politics here is complicated when science is invoked inappropriately for advocacy purposes.

Government *performs* science directly through its employees in federal laboratories and elsewhere. It seems obvious to me that this science has to be conducted with the necessary integrity, or it is a waste of public funds.

Finally, government passes and enforces laws regulating the conduct of science. It does so for reasons of public safety, national security, and ethics. Everyone agrees that embezzlement is not an appropriate way to fund a research project, nor is tormenting human or animal subjects an acceptable way to gain scientific knowledge. Governments regulate hazardous chemicals, and rare materials and key concepts for weapons of mass destruction. Most of these are not controversial, but some have generated a huge amount of publicity in the recent past.

I have focused on what you might call theoretical aspects of "Scientific Integrity in Government" because I do not think these issues are either simple or intuitive. Government officials operate in a highly contentious and at the same time highly regulated environment that leads to actions that are easily misinterpreted. We are dealing with the data of human behavior here, and we need to approach that data with the same caution that we use in our scientific work.

Thank you for inviting me to speak to this topic.