Minutes[*]

Faculty Consultative Committee
Thursday, October 31, 2013
1:00 – 3:00
Room 238A Morrill Hall

Present: Will Durfee (chair), Linda Bearinger, Avner Ben-Ner, James Cloyd, Eva von Dassow,
                Jigna Desai, Janet Ericksen, Gary Gardner, Maria Gini, Joseph Konstan, Russell Luepker,
                Alon McCormick, Karen Mesce, Ned Patterson, Paul Ranelli, Rebecca Ropers-Huilman,
                David Satin, Chris Uggen, Jean Wyman

Absent: none

Guests: President Eric Kaler, Provost Karen Hanson

Others: Deb Cran (Office of the Provost), Amy Phenix (Chief of Staff), Jon Steadland (Office of
                the President), Ken Savary (Office of the Board of Regents)

[In these minutes: the intellectual future of the University: quality]

The Intellectual Future of the University

        Professor Durfee convened the meeting at 1:00, welcomed President Kaler and Provost Hanson,
and provided a handout outlining the foci of the discussion (between the * * *). This is a special meeting,
held once per year, with the president and provost, to discuss issues related to the "intellectual future of
the University," and the topic today is quality. "Quality" is a term often used in discussing the excellence
of a university, but when one drills down, it becomes clear that quality has many facets, so it is
appropriate for the Committee to think about what it means. It is also timely because the strategic-
planning process will consider quality and the allocation of resources, and the president has referred a
number of times to excellence when talking about strategic planning; if one equates excellence with
quality, it is an important topic.

* * *

   Context
   • There is no absolute quality that can be used to measure all things for all people. Quality depends
      on values and context of the beholder.
   • Local norms determine agreed-upon definitions of quality, therefore global definitions are
      problematic
   • Quality is a matter of judgment, which must be exercised by a person, and cannot be externalized
      from the subject who judges.

   Purpose
   • Before defining quality, should we determine why we want to measure quality? Should the
      objectives drive the definition?

- Is quality about the output, or about the difference between input and output?

Metrics
- What is the (proper) role of reputation in defining quality?
- Is efficiency a part of quality?
- Do metrics drive behavior . . . in a good way? For example, with h-index and SRT scores as metrics, do faculty game the system? Or, do quality measures help or hinder interdisciplinary activities?
- If what our undergraduate and graduate students do after graduation is part of measuring quality, what should be measured?
- Can all of quality be quantified in measurable metrics?
- Time to degree (undergrad and grad) is an increasingly important metric. What should be its relative importance in determining quality?
- Are the metrics for determining absolute quality the same as the metrics for determining what should be improved?
- Quality definitions can be used for internal purposes (e.g. improvement) or can be used for comparison purposes (e.g. rankings). Can one set of definitions serve both purposes?

Granularity
- Does the quality of an institution equal the sum of the quality of its programs?
- We tend to look at program quality in isolation. Do programs stand alone or at a university are they inherently linked to their neighbors?
- Metrics involve averaging, which can smooth out peaks and valleys. Is that a good thing?
- How similar (or dissimilar) are the definitions for quality across the UMN campuses? How does campus quality roll up to UMN system quality?

External Forces
- Quality is to some extent being driven by external forces. Is that a good thing?
- How much freedom do we (UMN) have to determine our own definition of quality?

* * *

In the course of the 90-minute discussion, Committee members, President Kaler, and Provost Hanson made a number of points.

--   What does "quality" mean when it is applied to the academic missions of a Research 1 university? If one definition of its mission is to address societal issues for the long term and to educate students for their third or fifth job, not many research universities have communicated that mission well. The University must do so to differentiate its role from other institutional missions.

--   There is a relationship between quality and impact and also quality and efficiency. What if there is quality but nobody notices? Quality should mean something changes through education or scholarship. Efficiency to be included as well because it would be easy to have quality and impact with unlimited resources. The University should not neglect strategic advantage even where it does not start out at the top, it should improve students who are not at the top, and needs to make the most of the people it has in research in producing an impact—it should measure what it can do with the resources it has.

--   The previous comment repeatedly referred to change. However, the mission of a university also encompasses conservation of knowledge and continuity in its cultivation. The term "quality" has no content: there is no quality, in the abstract, but only qualities—decrepitude, elegance, efficiency, and so on ad infinitum. The discussion should address what qualities are appropriate to a research university, rather than seeking to paste metrics onto the empty abstraction "quality." One appropriate quality has been noted already, namely, that a university—unlike almost any other kind of present-day institution—works for the long term. Its role is not merely to respond to ephemeral market demands but to carry out education and research that endure over time.

--   The University exists in a turbulent environment and decisions cannot always be postponed. The University will have to re-evaluate what counts for scholarship, what it needs to do other than reproduce itself. Interdisciplinarity is very complicated; who is to be hired today? Do departments hire in contested areas or hire within existing fields? Do they take into account shifting demand? Does the University give society the medicine it believes is necessary and useful for society, even though society may not want it?

--   This is not a new debate; medicine has been wrestling with it for a decade and has learned its lessons. One, many of the metric questions on this document have been answered in the literature; they were hotly contested, resisted, debated, and are now settled and accepted as reality. This same process will happen across the University (there will be publicly-reported metrics, financially-tagged metrics, that have to demonstrate what is being produced). Two, there are experts in this field and for at least some questions, there are right and wrong answers. Medicare and Medicaid commissioned a process that was extremely extensive and the best in the world; it had teams of experts involved. There are many kinds of metrics and there are experts around the University on them. Additionally, there can be incentives tied to metrics, and they do drive behavior (good and bad), but this is well-trod ground in medicine. It is happening so the University must get on the train. The University should determine if that is a good thing and how we respond, but this is the reality. [The fact that it is happening everywhere doesn't mean the University should do it. Rather, it is a university's job to determine whether or not it is a good thing to do. Metrics analogous to those applied in medicine are not appropriate to many areas of academic work; it is unsound to paste metrics and words onto an ill-defined notion of "quality."]

--   The history of higher education is replete with dichotomies of values where both have been supported at various times: exclusivity versus inclusivity, usefulness of scholarly work versus scholarly work that is considered removed from society, students satisfied versus simply surviving, retention and graduation versus weeding out. Even today people hold to different poles on certain values. This Committee has to think about the whole University, and when quality is being measured, it is sometimes hard to recognize and hold all values at the same time. In one contemporary example, our current graduate program reviews suggest that a higher doctoral graduation rate is better. But what is the ideal graduation rate in doctoral programs? 30% 50% 100%? What level of inclusivity or exclusivity suggests quality? There are different values on that one question all over the University.

--   The train has left the station as far as metrics go and the University will increasingly see metrics from the state and through student choice. But metrics are dangerous, especially if they do not measure what is actually valued, but instead what is easiest to measure. For example, when the University adopted 4-year and 5-year graduation rates as a metrics of great importance, what it also said was that it did not welcome part-time students, students who may have to work full-time while pursuing a bachelor's degree. And it seems that it is now doing the same thing with graduate programs, which is worrisome because it cuts off an important constituency and part of the mission. Overdependence on the wrong metrics also

leads to decisions made through measurement rather than disciplined thought, so (for example) programs trying to optimize their graduation rates may only admit "safe" students who cannot help but succeed while turning away riskier but potentially brilliant students—ones who could have a huge impact (not measured) but might not graduate (measured).

--   People in the University are all over the map on fundamental questions; when one thinks about what the University wants to be, one must recall that it is set in a larger social context. It has a role in constructing society but must also respond to it, and often the University does not think enough when it must do planning for how it fits in society—and not everyone will end up with the same view on that question.

--   The question is whether one can meaningfully measure that which one believes matters. The history of medicine once included as the dominant view, that one cannot apply statistics to something as complicated as medicine. Similar arguments resurfaced a decade ago about quality measures. The process is still in its infancy, and developing new statistical methods such as risk adjustment factors. The University will have trouble accurately measuring the quality of what it does but that does not mean there is not a way to do it—taking a quantitative approach to quality improvement is, across so many fields, the next generation's way forward. But it is hard to do and the process is just beginning. [Can one seriously imagine it possible to "measure" the quality of a historical explanation, or a linguistic theory, or a literary interpretation? Where what counts is what one can count, one should do so. But where that isn't so, the use of metrics becomes a way to avoid the responsibility of exercising judgment.] It is absolutely true that metrics can judge the arts; there are things in every field that one can say are good, but there are ways the process can go wrong. But it is possible to say what is good and bad in every field. [By means of numbers?] It is; people measure overtly or they measure covertly.

--   On the question of difficulty of measurement: quality is a latent construct with multiple measures. Qualities of communities are often measured this way, and while the individual measures are not always great, they can be combined to get at questions of quality. One has to start with a vision: what is a great department or program? The measures might include impact, visibility, leadership, satisfaction. While some factors are more easily measured, there are ineffable qualities as well. One does not want to lean too heavily on the easily quantifiable, bad measures, but there is often surprising inter-rater reliability on many of the basic quality assessments. One can be "pro metrics" while responding to differences and not using the metrics mindlessly.

--   If it is possible to use metrics for clinical care, there are similar metrics for other activities; for undergraduate education, metrics for the University will be similar to those used by MnSCU. What about metrics that measure long-term impact? Such as Ph.D. placements of people who later become leaders—the University must look at factors that distinguish it from other Minnesota institutions.

--   In thinking about our behaviors—what to do—in a financially scarce era, the faculty knows we have difficult decisions due to less money. The question of how to distribute it drives the discussions of metrics. The assumption is that metrics will determine the best allocation for funding. Once we know what is good and can measure goodness, we will know what to invest in. But we don't have a single definition of goodness; this institution has many definitions of goodness. We do need to look beyond measurements to shared values. Values rather than metrics should drive our decisions. As academics, when we make judgments and arguments, we go beyond quantitative proof; we also use qualitative evidence to support our judgments. Not everything can be measured. And we need to think in terms of

values, arguments, and evidence, not just measurements/metrics. Some of our shared values should be how we think about society and how we invest in the common good (like the achievement gap). As a collective we need to address and invest in this.

--   The definition of measures and metrics does not mean there is no human judgment. For example, to evaluate a restaurant, one may depend on aggregating the opinions of a collection of humans to form a metric (e.g., a number of stars). It is also important to recognize that not all metrics are means or medians; metrics can identify outliers with interesting properties. For the University, a department may explicitly want to hire a faculty member that 95% of peers think is a bit of a crackpot but that 5% believe is a world-changing future Nobel laureate. Taking chances on people and programs that are not popular, but where some trusted individuals strongly believe they are outstanding, is part of how fields change and grow.

--   The discussions about the University's aspirations are recurrent and rarely reach closure. President Bruininks proposed that the University aspire to be the #3 public research university but never defined the metric or implemented a plan to reach this goal.

        The University is now asking the same questions. One hopes that the strategic planning group will define its aspirations and provide a general plan of how it attains those goals. Excellence in the University's teaching, research and service should be the driving force and the metrics of success or failure should be objective and recognized by external bodies.

        Several other considerations should be included:
        1. Funding: The 'new normal' is here and real. There will be no increases in funding unless the University is creative in finding new sources. It cannot continue to increase tuition, and research funding will be more challenging.
        2. Scope: It cannot continue to grow more new programs if it is unwilling to close others. The University does too many things already and some units need to go.
        3. Mediocrity: The University cannot be the state's leading institution and tolerate OK programs and OK faculty. It is stifling intellectually and undermines the mission.
        4. Leadership: Leadership comes from the top and faculty governance needs to get behind the coming strategic plan.

--   The discussion topic could really be "the future of the University." It must deal with a bucket of money that will not get larger. If it thinks about ranking, it will be asked if it is the best. The University must do well what it does and must be careful what it commits to. It must commit to do well what it does, and in some cases what it does will be #1. It is necessary to distinguish between quality and rank-order; if the University can do something well but it is not #1 in the field, that is OK.

--   The train has indeed left the station on metrics. I think the University will get metrics thrust on us, as was the case for Texas A&M, if we do not develop our own.

        43,000 students applied to the Twin Cities campus (which is a metric), the University has metrics from the legislature, which we will meet

        It is indeed possible to parse quality in subjective fields, even if the metrics are not ordinal; departments make such decisions every time they hire a faculty member. The conversation around

"quality" and "qualities" is interesting, but one can make measurements in sociology and economics, for example, and can begin to reach conclusions (e.g., on the achievement gap, which schools are high and which are low in measures to close the gap). There is subjectivity and judgment that goes into the decisions, and that is something that the faculty must do as a group; there are measures that doctors, poets, engineers, and musicians use. An engineer can't judge music, but others can and they must do so.

Resource allocation requires judgment and the institution should not invest in programs that do not improve student outcomes. The University must engage in the conversation about quality ourselves because we do not want others to do it.

--   Just because a judgment is subjective does not mean it is not reasoned. Academics do ratings and make judgments every day. They do so in spades, but people are more uncomfortable doing it at the institutional level.

--   A more specific issue related to quality that the University faces is this: as it moves more into online education, some hybrid courses and some entirely online, it will be possible to have outstanding faculty in them and get students through. How will the University measure the quality of those programs? It is not apparent that there are good metrics for them.

--   The term "measure" is being confused with "evaluate." There is no meaningful measurement of a poem's quality. It is true that subjective judgment involves reasoning (not measuring). Turning to the use of judgments about quality to allocate resources, suppose a program is judged to be low quality: how is that evaluation to be acted upon? Does it mean the program gets cut? Suppose the fields it covers are integral to the university's mission; does the program instead get more resources so as to improve it? The term "quality metric" is an oxymoron. What is the purpose of proposing to measure quality -- to identify criteria for cutting programs?

--   There is a need for continuity and to keep knowledge intact. The University cannot abandon knowledge essential to humanity, even if it is not producing cars. But the institution must set priorities for the next 10-25 years, so what mix of faculty does it want in terms of continuation versus pushing ahead where there is no industry? In departments, across the University? One would think that the University would want the 5% who create new fields. The institution must take responsibility for making decisions about how to produce various kinds of knowledge and what institutional support will be required.

--   One must look at interdependence and centrality. On the matter of quality, evaluation, and metrics, take a hypothetical department of plastic bottle labels. If it is ranked #1 in the world, the University should retain it, even if some do not see value in the field at all, because it is worthwhile to have things that are great—but the University would eliminate it if it were ranked 35th out of 36 such departments. If a department is one that the University cannot do without—one that is central to an undergraduate education, depended upon by many other departments, or perhaps the only program of its type in the state or region and highly valued by the state—there must be investments made to make it better. The faculty can and should identify the criteria for making these decisions; but the process requires a leader who will actually make the decision to shrink or eliminate a program—or the entire strategic-planning effort is a waste of time.

--   To what extent are traditional internal measures of quality in any particular field now overwhelmed by external pressures, such as perceptions of relevance or economics?

--   If one considers two principal drivers of investment decisions to be quality and centrality, where both in a program are high, the decision is easy. Where neither is high, the University needs to disinvest. For example, no Research 1 university will be without a Department of Economics, but it can have a medium-quality department and be OK. Where possible, it is necessary to decouple temporary variations in undergraduate demand and long-term scholarship in a field. Enrollment in fields will have fluctuations and the institution must manage the fluctuations if the graduate program and research productivity is excellent. It is important to recognize different scales of time.

--   In the field of adolescent health, one can predict with a high degree of accuracy how well an adolescent will do by using just four questions: Do you like school? Do you feel your teachers care about you? How much do you feel your parents care about you? Do you have an adult you can talk to about problems or concerns?   Because their predictive capacity, these questions are used one-on-one in clinical practice and they are assessed and compared at community levels, state levels, and compared globally, country to country. Few doubt that anyone would argue that asking teenagers these questions assesses something that is subjective, yet they lead to quantifiable measures that are highly useful for understanding and predicting outcome.   This example reaffirms what others have been saying today, i.e., that researchers do have the capacity to measure subjective characteristics and qualities.

--   All in the academic realm are good at critiquing and evaluating. They do it constantly—they grade papers, serve on study sections, judge meritorious work of colleagues, and vote on promotion and tenure. So they should be able to create methods and measures that will guide future priorities and decisions for our institution.   It's what they do in their faculty positions and they are good at it.

--   In terms of the interdisciplinary questions on the list, most of which are under the heading, "Granularity," one can fear that the focus may be on how to successfully evaluate quality at the school, discipline, or departmental level, but stop short of creating strategies for measuring quality in interdisciplinary endeavors. Sometime in the past few months an article was circulated to the Senate Research Committee that ranked universities based on characteristics related to faculty publications. The University of Minnesota was ranked in the top 10, possibly #5, in terms of the collaborative nature of its publications. This is a terrific quality for which to receive a high ranking; it reflects the institution's interdisciplinary capacity. But, if the University does not create measures for assessing interdisciplinary qualities, it would miss this sort of uniqueness.   With regard to one specific question on the list—is the quality of what the faculty do interdisciplinarily greater than the sum of its parts, i.e., individual programs—the answer is yes, and the institution does not want to miss that as it develops metrics for measuring quality.

--   Related to the issue of what qualities are to be measured, the University must first ask itself what it is that it values. As others have pointed out, certain characteristics do not lie on a single continuum, e.g., admitting those from the top 10% of high school seniors versus admitting those who may be the first in their family to go to college but who may not be in the top of their class. These should be measured as two separate qualities, if the University values both.

--   It is necessary to look also at the quality of the University system and to be sure it is being used to the highest level possible. That is not happening now. [The president expressed agreement and said the

longer-term goal is to get activities working together and leverage programs; once the Twin Cities strategic planning is completed, he will establish a group to identify areas of synergy across the campuses to help the institution be better as a system.]

--   The University needs to do the right thing right at the right time. Quality can also mean lack of defects or flaws. Part of the definition of quality is consistency, which can battle with innovation; institutions evolve because there are mutations.

--   One concern is that metrics are numbers, properties that allow ranking. It is easy to determine if a program is #1, in which case the University keeps it. If a program is #35 or #36, it may drop it. What is more difficult is those that are in the middle area, where the rankings are muddled and not significantly different. All agree on who the spectacular faculty members are and who the bad ones are; what is tricky is evaluating the ones in between.

--   One risk in the discussion about quality IS centrality: there is a striking difference between many of the doctors who are upset by performance measures versus those who welcome the monthly reports. Those opposed frequently see themselves as peripheral and disposable to their parent organization, such that a bad report this month could be their last. In contrast, those who feel central to their organization see a bad monthly report as helpful evidence that they need more help and resources to succeed. Because the perception of one's centrality to an organization is so critical to one's world view of quality measurement, it is practically difficult to separate the two. In practice, in order to settle questions of quality measurement, the question of centrality must be settled.

--   The amount of money the University is likely to have will not grow dramatically in the near term but strategic planning is likely to produce a number of good ideas. For the University to continue to do everything it is doing would not be wise, but at the end of the day the faculty own the conversation about how programs stack up and how they can be merged or leveraged to be better.

--   With respect to the difference between measuring and evaluating, creativity is important. It is not the number of papers, it is their novelty and potential importance to society that matters, and that is difficult to describe to outsiders. The University needs to describe creativity in a way that the public can understand.

--   In terms of the #1-ranked bottle-labeling department, Professor Gardner met, while in graduate school, the heir to Dietzgen, the best slide-rule company in the U.S. The University must make judgments about what may not be the future.

--   In terms of limited dollars, that means the University must give up something as it moves into new areas. One opportunity to do so is the device of presidential initiatives and programs like MnDRIVE. A few key positions can leverage others to change direction.

--   It is important to measure the accomplishments and achievements of the alumni, both early and later in their careers. One can then ask if the University is doing as well as its peers.

--   If the University has a water-conservation department, it might not invest more in the plastic bottle label department: it is not just centrality and excellence, it is a question of what institution's goals

and values are. When looking at what is best, there may be eternal departments, but it is necessary to look at what each is doing and have conversations across departments.

--   One should argue for both humility and boldness. With respect to the department of slide rules, at one point studying electric grids (in electrical engineering) was old-fashioned, but now experts in that field are in high demand. One must have a sense of humility: sometimes an institution will keep a field around because those who argue it is obsolete could be wrong. If one thinks of the University as a portfolio in which to invest, to be successful it must have programs that are successful at different times. For instance, if the University has a large component related to clinical medicine, it should have other components that will succeed at times when medical revenue may be down. Also, one would not put all one's money in bonds; there should also be some small cap funds. So there will be high-risk small programs, some of which will become obsolete and some of which will be immense successes. And even within departments, the University wants a mix of "dependable" faculty who will succeed, but may not be stars, with investments in potential stars who are a larger risk, but with a larger potential reward. This is a challenge to the University, the colleges, and departments: they need to take chances so they do not end up under-performing their peers because the University was not bold.

--   The high-risk, high-reward programs may not generate much indirect-cost money but colleges do need them in a balanced portfolio.

--   The University needs to diversify its offerings to ensure that it has programs that meet the needs of the institution's multiple stakeholders. It needs to have some programs that represent the best in innovation and might be risky in terms of their reward. It also needs to have some programs that represent ongoing, long-standing relationships with its various communities. As it refashions itself for the future, it needs to think carefully about how to ensure the University remains connected to its various stakeholder groups.

--   One appreciates the point that creativity is not measurable while also understanding the pragmatic need to determine how to allocate resources. The use of metrics to manage resource allocation may drive out the very qualities the University claims to value (the School of Music just lost a composer, for example). There is a conflict between cultivating qualities such as creativity and imagination and managing programs by assessing quality in quantitative terms.

--   As the University determines how to measure itself, it may want to look to the outside world where the practice of measuring, particularly measuring what people think, has been taken to a whole new level. For example, the Amazon.com recommender system aggregates the thoughts of people like you when it suggests the next book for you to read; with surprising accuracy. When the University measures attitudes of recent graduates, it is still stuck in the 5-point Likert scale questionnaire model and the exit interview. Why not instead have a Wiki-like page that recent alums could openly edit, and like Wikipedia it will gradually reach an equilibrium that represents a consensus of how alums view their experience at the University? As we think about how to measure ourselves, let's assess how the world measures itself now, how it will measure itself in ten years, and use some of those methods.

--   What also comes up is scale: there are investments in big programs, but there are a number of small programs and projects the University could invest in, and they do not always take a large amount of money. Often the discussion is about large programs but there may be small ones worthy as well, not requiring millions of dollars—perhaps one faculty position would make a difference.

Professor Durfee thanked everyone for a stimulating conversation and said that it will feed into the strategic-planning process. He adjourned the meeting at 2:45.

-- Gary Engstrand

University of Minnesota