**Proposition: 16-18/19**

**Title:** A proposal to change the current NMSU course evaluation to remove the collection of all quantitative, letter, and numerical items from student course evaluations.

**Date Submitted:** 22 Feb 2019

**Sponsor(s):** R.T.J. McAteer (A&S), M. Hamzeh (A&S), L. Keleher (A&S), A. Orzoff (A&S), R. Conley (NTT), M. Prentice (Edu), B. Corran (DACC), J. Parra (Edu), E. Kuchler (HSS), S. Beck (Lib), C. Voisine (A&S), S. Stochaj (Eng), C. Brown (A&S)

**Proposed Committee:** Faculty Affairs

**Assigned Committee:**

**Prior Approvals:** None

**Proposal:**

NMSU should only collect student course evaluation using qualitative or ranking data, and accordingly revise NMSU ARP as needed. As no part of existing NMSU ARP calls specifically for a quantitative, letter, or numerical items to be included in student evaluations, and as the state law also does not call for such items to be included in student evaluations, a change to NMSU policy may only require a statement in ARP 9.20 Part 2, new section I (or clarifying statements) such as

"[I] Student evaluations considered as part of performance evaluations, may not include numerical ratings, letter grades, or other "scores" of specific aspects of the course, the course as a whole, or of the instructor"

The NMSU faculty senate requests for Legal Counsel to alter the NMSU ARP, and commits to working with all stakeholders, including the Provost’s office, Legal Counsel, Deans of Colleges, Instruction Innovation and Quality, and the Teaching Academy, to implement these changes throughout the NMSU system in a timely fashion.

**Rationale:**

While student evaluations of individual courses are one critical component in ensuring courses result in the best student learning outcomes and continuous course improvement, mounting evidence shows that the collection of quantitative, letter, or numerical items as part of student
course evaluations discriminates against women, people of color, and other marginalized groups of faculty, impedes course redesign by faculty, and removes the student from critical evaluation of their own learning.

State law NMSA 1978 Section 21-1-7.1 (see Appendix A) mandates that “The boards of regents shall ensure that a peer review is afforded the faculty member and that student evaluations are considered in the evaluation of the tenured faculty member's teaching.” As such, we have a pedagogical and legal obligation to ensure student evaluations are collected and considered in the most effective and meaningful manner. Quantitative, letter or numerical items as a component of student evaluations impede this effort and should hence be removed. The reasons for this impediment are threefold

1) **Numerical student evaluations consistently discriminate against women, people of color, and other marginalized groups of faculty.** Several recent research articles (e.g., [1,2,3]) show that quantitative, numerical, or letter items in student evaluations measure student biases better than they measure teaching effectiveness. As such, inclusion of these types of items in student evaluations inhibits the University's more meaningful and equitable processes of tenure and promotion, and hence supports another structural barrier for women, people of color, and other marginalized groups of faculty. This is directly contradictory to NMSU’s position as a federal-funded and Hispanic-serving institution.

2) **The collection of numerical student evaluations may impede changes to course design and discourages personal reflection on teaching.** Professors who obtain 'high' scores on numerical student evaluations are not prone to incorporate course redesign and are open to accusation of grade inflation as a quid-pro-quo. Professors who obtain 'low' grades in numerical student evaluations are prone to play it safe and only incorporate the simplest course redesign methods, focused on obtaining better evaluation scores [4]. In neither case is course redesign, and changes in teaching approaches, focused on improved student learning [5, 6]. Often the most effective teaching methods can lead to lower evaluations in short run as the teaching methods that may be most effective for student learning are challenging to students and thus garner low numerical evaluations.

3) **The collection of purely numerical student evaluations does not require students to critically evaluate the course or take an active role in the teaching and learning process.** Numerical rankings provide a simple method to include student evaluations, without probing for effective feedback from the student as active agents in their own learning and the overall teaching process. Well-designed qualitative and ranking questions gather much more effective student feedback and create an environment in which the students assess their own goals and effort in the class. From this feedback, a professor can make those changes, as required, to improve student learning [7, 8].

Rather than inaccurate numbers on how much teaching should be enhanced, faculty would obtain much better feedback on what should be enhanced. This can be best provided by qualitative data and leads to continuous improvement of student learning. The specific nature
of the inclusion of student evaluations will continue to be determined by each college and department, as is current practice. Such inclusion could include best practices other than numerical Likert-scale questions, such as an end-of-semester survey to include questions that rank various criteria of teaching effectiveness. One such survey, adapted from the University of Oregon, appears on the Teaching Academy website at https://teaching.nmsu.edu/teaching-effectiveness/ and is Appendix B to this proposal. The results of this survey could feed directly into two other measures of evaluation—“evidence from the instructor” and “evidence from other professionals”. Once an instructor gathered information about the area of their teaching that represented the students’ perspectives on their greatest strength as well as the greatest challenge, the instructor could self-reflect on how to build on their strength and address their challenge—or they could ask a peer to give ideas by reviewing their syllabus or attending class [9].


Available at https://doi.org/10.1007/s11199-014-9317-4

Available at https://doi.org/10.1007/s11199-017-0872-6

Available at https://doi.org/10.1016/j.stueduc.2016.08.007.

DOI: 10.14293/S2199-1006.1.SOR-EDU.AETBZC.v1
Available at https://www.scienceopen.com/document?vid=818d8eco-c908-47d8-86b4-5dc38f04b23e

Available at https://www.chronicle.com/article/Students-Evaluating-Teachers/245169?fclid=1wAR3o-ZkF8lQemPrwauBpUO2w3uRRVT3qWkw/9tNiW-iLYncXMP-MD8RnCamU

Available at https://www.chronicle.com/article/Colleges-Are-Getting-Smarter/245457


21-1-7.1. Post-tenure review process required.

A. The boards of regents at all state baccalaureate degree-granting educational institutions are authorized to direct the president of the university to institute a periodic post-tenure review process for all tenured faculty.

B. The boards of regents are authorized to direct the president of the university to establish programs designed to assist faculty members in enhancing their teaching skills.

C. The tenured faculty member shall be subject to review every three to five years based on a review of a number of factors, including the following:

1. an evaluation of the faculty member's teaching;
2. an evaluation of the faculty member's research and scholarly output; and
3. an evaluation of the contributions made by the faculty member in the area of public service to the institution and the community.

D. The boards of regents shall ensure that a peer review is afforded the faculty member and that student evaluations are considered in the evaluation of the tenured faculty member's teaching.

E. In the event a faculty member receives an unfavorable evaluation in the area of the faculty member's teaching, the post-tenure review process shall include:

1. a two-year probation and reevaluation period; and
2. loss of tenure if, during the subsequent probation and reevaluation period, the faculty member fails to demonstrate improvement in the area of teaching.


Emergency clauses. — Laws 1995, ch. 150, § 4 contained an emergency clause and was approved April 5, 1995.

Temporary provisions. — Laws 1995, ch. 150, § 3, effective April 5, 1995, provided that the board of regents of all state baccalaureate degree-granting educational institutions shall study the options, advantages and disadvantages of developing a procedure for granting tenure based solely on a faculty member's teaching ability.
Note to Faculty 1: The following instrument was adapted from the Student Experience Survey developed by the University of Oregon at: https://provost.uoregon.edu/revising-uos-teaching-evaluations/. The characteristics identified as essential to teaching come from the Oregon definition of excellent teaching at: https://tep.uoregon.edu/teaching-excellence. Used by permission.

Note to Faculty 2: In order to get sufficient feedback to be helpful, you may want to reiterate the material in the preamble below when you administer the evaluation.

End-of-Semester Student Evaluation

Feedback is important for instructors just as it is important for students. To be most helpful, feedback should be specific, offer examples, and be actionable; that is, something that the teacher can take action on. For example:

<table>
<thead>
<tr>
<th>Instead of...</th>
<th>This is more helpful...</th>
</tr>
</thead>
<tbody>
<tr>
<td>“This teacher gives unclear instructions.”</td>
<td>“This teacher gives unclear instructions. For example, we did not understand the length or the detail that was required on our final paper. A successful sample paper would have helped a lot.”</td>
</tr>
<tr>
<td>[This statement leaves the instructor wondering what is unclear about the instructions and how to make them clearer.]</td>
<td></td>
</tr>
</tbody>
</table>

☐ I understand that my responses will be anonymous, and that my instructor, the course department head, and other institutional personnel (college dean, promotion committees, etc.) will also be able to read my anonymous responses after the final course grades have been assigned.

1. Describe what most helped you learn in this course and what most interfered with your learning—and why. Please write one full paragraph so that your teacher understands your experience in the course.

Continued on next page
The following have been identified by the [Department/College of ___________________________] as essential characteristics for excellent teaching: *(Note: departments/colleges may want to revise this list of characteristics based on their own definitions of excellent teaching.)*

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Associated Student Experience</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Inclusiveness</td>
<td>I felt welcome and respected.</td>
</tr>
<tr>
<td>(b) Transparency of instructions &amp; grading</td>
<td>I understood instructions for assignments and how I was graded.</td>
</tr>
<tr>
<td>(c) Feedback</td>
<td>I received sufficient feedback on my work in an accessible format and in time for me to make changes and improvements.</td>
</tr>
<tr>
<td>(d) Challenge of the course</td>
<td>I felt challenged and I learned a lot in this course.</td>
</tr>
<tr>
<td>(e) Quality and accessibility of course materials</td>
<td>The course content, materials and assessments (texts, notes, slides, videos, exams, projects etc.) were high quality, accessible, and helped me to learn.</td>
</tr>
<tr>
<td>(f) Instructor Support</td>
<td>I felt the instructor wanted me to succeed in this course.</td>
</tr>
<tr>
<td>(g) Engagement in out-of-class assignments or projects</td>
<td>The assignments or projects regularly increased my interest in the course.</td>
</tr>
<tr>
<td>(h) Active Learning</td>
<td>The course activities actively engaged me in my learning.</td>
</tr>
<tr>
<td>(i) Quality of the interactions between students</td>
<td>I had opportunities for high quality and collaborative interactions with other students.</td>
</tr>
<tr>
<td>(j) Instructor communication</td>
<td>Communication with the instructor was positive, timely and meaningful.</td>
</tr>
<tr>
<td>(h) Other</td>
<td>(describe)</td>
</tr>
</tbody>
</table>

2. **Greatest strength:** Which **one** of the essential characteristics listed above was most beneficial to your learning?

   Explain specifically **how** this characteristic of the course helped your learning.

3. **Needs Improvement:** Which **one** of the essential characteristics listed above was least beneficial to your learning?

   Explain specifically **how** it could be improved to help your learning.
4. **What else you would like to say?**

5. **Thank you** for completing this survey for your course! This is a beta version of a new way to get helpful feedback from students to instructors and departments for continual course improvement. The following two questions are for the task force that built this survey (not your instructor). What did you think of this new survey as a tool to provide feedback to your instructor about your experience in the course?

- Great
- Good
- Fine
- Bad
- Really bad

Do you have any suggestion for how this survey could be better? (e.g.: specific wording, format, questions etc.)? If so, please give your suggestions here:
“Teaching Eval Shake-Up”
Recent articles in Inside Higher Ed (“Teaching Eval Shake-Up”) and in the Chronicle\(^1\) called attention to problems with teaching evaluations, including gender bias. Many in higher education are questioning how, or even whether, SETs should be used in the evaluation of teaching effectiveness, i.e., as part of annual performance evaluations or promotion and tenure decisions. The University of Southern California and the University of Oregon have recently adopted sweeping changes to their teaching evaluation processes. These schools have developed many alternate teaching effectiveness evaluation materials available at http://cet.usc.edu/ and https://provost.uoregon.edu/revising-uos-teaching-evaluations/.

This document was created to support discussions on meaningful and equitable evaluation of teaching effectiveness. It includes brief descriptions of possible alternatives for evaluating teaching using Likert-scale SETs and evidence from the instructor and other professionals (i.e. peers). It provides a summary of resources available from the Teaching Academy.

Triangulation
There is widespread agreement that effective teaching evaluation requires the collection of data from at least two sources, i.e., triangulation. University Administrative Rules and Procedures (ARP 9.31) encourages the triangulation of data for the evaluation of teaching effectiveness using evidence (a) from students, (b) from other professionals, (c) from the instructor, and (d) of student learning. Triangulation has advantages because there are no perfect measures of teaching effectiveness. University of Southern California now relies mostly on peer review instead of student evaluations. University of Oregon now relies on rankings of criteria for teaching excellence coupled with self-evaluation by the instructor.

Evidence from Students
State law requires that student evaluations be used as one form of evidence in the evaluation of teaching. Given that students have an important perspective on our teaching, how should their feedback contribute? One promising possibility is being explored by the University of Oregon. In the online evaluation that they use, students rank the area of greatest benefit to their learning and then are asked to explain how it helped their learning. Students then rank the area of least benefit to their learning and explain how it could be enhanced. An adapted version of this Oregon evaluation can be found on the Teaching Academy website at

\(^1\) https://www.chronicle.com/article/Colleges-Are-Getting-Smarter/245457?key=nbhuwAvuzuO_LrP40ugFwBkgO0W35m70QQw43EnNlhbh6MoY4OiwQlpbf2KOryQaWpWTWkJR3pmUW14NiU3NksqUXg4U1ipSW3lQTXYMWs1UkljZHBvSiFOYw&fbclid=IwAR3H23zZKRIBx6YaMLEzpiHwTMNUsDXhrvKdsGk4wslf9y9NdGw1P7PHw
Evidence from the Instructor and from Other Professionals

Ranking data from students provides important information to instructors and their colleagues, which informs “evidence from the instructor” and “evidence from other professionals.” For example, assume “instructor communication” was ranked as the greatest benefit to student learning with reasons given as to how that benefited students. The instructor and other professionals could then reflect on what could be done to further build on this benefit. The same process could be followed for the area of least benefit to student learning. Every teacher, no matter how strong or challenged, would be encouraged to address both their greatest benefit to learning and their least benefit, using self-reflection as well as course observations of face-to-face classrooms or online courses. This would further instructor continuous improvement and lifelong learning.

https://teaching.nmsu.edu/teaching-effectiveness/). [Note: this instrument is not yet posted. . . it is attached.]