Recommendation of the Provost’s Task Force on Student Teaching Evaluations (STEs)

Task Force: Larry Overzet (chair), Abhijit Biswas, Gail Breen, Harold Clark, Peter Ingrao, Simon Kāne, Kim Knight, Angela Mcnulty, Sai Nikitha Prattipati (student), Paul Diehl (ex officio), Karen Huxtable-Jester (ex officio)

Task Force Charges:
1. Review the format, questions, and conditions of administration for student teaching evaluations and make recommendations for any changes therein.
2. Construct a short set of guidelines to assist decision makers in interpreting and using student teaching evaluations.

Recommendation 1: Rebrand the Experience
Reframe and rebrand the “evaluation” process and method to emphasize “Assessment of Learning Experience (ALE)” (instead of “course evaluations”, “instructor evaluations”, “end-of-term evaluations”, etc.).

Overview & Reasoning
Remove the negative association with the term “evaluations” for both faculty and students. Faculty have frequently expressed concern that students are not qualified to “evaluate” their performance and therefore may avoid evaluations and do not promote or support the process. Students may be unclear about what they are evaluating or may have misconceptions about what “evaluation” means. Reframing the process as an “assessment” of the learning experience from the “student” perspective could allow the university to emphasize the positive aspects of the process, increase participation, and improve the quality of responses.

Recommendation 2: Standardize the Assessment Timing
Standardize the period of ALE availability to the last two weeks of class (i.e. two full weeks before final exams). Although not all classes will be able to conform to this schedule, the majority should use the allocated time window to perform assessments.

Overview & Reasoning
Improve the validity of quantitative responses by minimizing (as much as possible) variations in timing. Increase participation rates by developing a customary and expected window of time when assessments take place.

Recommendation 3: Standardize the Assessment Procedure
Develop and adopt a standardized and documented procedure for all schools/departments/units to use. Designate a specific 15 minute window of class time to present and execute the procedure.

Recommendation 4: Keep the Assessment Procedure Exclusively Online
Continue to perform the assessment process exclusively online.

Overview & Reasoning
The UT Dallas community has experienced some debate regarding impact of the mode of assessment (electronic vs. paper) on both the quantity and quality of results. Some faculty members have proposed a return to a paper-based system, while students prefer to use an online system.

Although not mandated by the UT System’s Recommendations for Incorporating Student and Peer Reviews in the Faculty Evaluation Process (https://dox.utdallas.edu/policy1214) – the 2012 task force “strongly recommended” that we “utilize an online system for course evaluations.” They reasoned that an “online system is more economical and sustainable than a paper-based system, providing quicker results and offering greater ability to perform data analytics.”

The committee agrees with this characterization of the benefits of an online system. It is recommended that UT Dallas continue to use an online system exclusively.
Recommendation 5: Add Guidance Information to ALE Results

Include both a cautionary statement ("black box warning") and comparative information with all assessment reports and data extracts. In addition, develop a more detailed set of guidelines on the proper use of ALE data. This can cover the material at further depth - include a link to this information with every report.

Cautionary Statement:
The cautionary statement is intended to briefly explain the limitations of the information presented in ALE results and how to interpret the information effectively. Include a link to detailed guidelines.

Comparative Information:
Comparative information should be designed to provide guidance on how the data relates to similarly assessed items (i.e. a yard stick to understand what the numbers might mean).

For example, the assessment numbers from a calculus section would include statistical information about other similar sections - this might show that a score of 3.6/5 for a calculus course is average whereas a score of 3.6/5 for an algebra course might be lower than the average.

Guidelines for Interpreting ALE Results:

- ALE results are more a reflection of the 'student experience' than a direct indicator of 'instructor effectiveness'.
- ALE scores and metrics may not be statistically valid and inferences made from these may be subject to large margins of error.
- ALE results are subject to errors due to small samples, low response rates, and self-selection biases (only students who choose themselves to respond).
- All student comments should be carefully read and general holistic and summative trends should be qualitatively identified.
- ALE results should be viewed in the larger context of the instructor’s historical results (where available) over a period of time and over multiple courses.

Overview & Reasoning
Historically, evaluation reports and data have been used in processes ranging from performance reviews, award allocation and university/school administration without understanding the limits of the information presented.

Providing some way to compare similar groups of data along with a better understanding of the limits of such comparisons could improve the quality of the decisions made using ALE data.

Including a cautionary statement with reports and data-extracts increases awareness of the limitations of working with assessment/evaluation data. A separate detailed set of guidelines provides the supporting material and justification for the warnings without clouding the core message.

Recommendation 6: Promote Understanding of Teaching Effectiveness

Update university practices and policies to adopt a more comprehensive and complete definition of teaching effectiveness.

Student assessments of faculty and courses should never be the exclusive or even primary method for evaluating faculty instructional performance. They should be considered as one element of one dimension (classroom performance) among many dimensions of instructional performance.

Consistent with this, the university should encourage a holistic review of teaching using the following guidelines:

- Promote the idea that instructor evaluation should broadly consider multiple pieces of evidence of performance. UT Dallas’ Center for Teaching and Learning (CTL) has produced the “Evidence of Teaching Effectiveness” manual which outlines best practices for evaluating teaching performance using multiple sources. (https://dox.utdallas.edu/manual1076)
- Provide reviewers involved in annual, probationary, tenure, promotion, award, and related reviews of instructional performance with the “Evidence of Teaching Effectiveness” manual during every review process.
- Provide reviewers with the “Evidence of Teaching Effectiveness” manual so they can better prepare review materials and not place too much emphasis on the wrong evidence.
- Review and update university documentation and policies to align with the ALE procedure and broader methods of measuring teaching effectiveness.

Potentially Impacted UT Dallas Policies:

- UTDPP1077: General Standards and Procedures Faculty Promotion Reappointment and Tenure
- UTDPP1089: Annual Review of Faculty
- UTDPP1062: General Standards and Procedures for Review of Nontenure-System Faculty
- UTDPP1064: Procedures Governing Periodic Performance Evaluation of Tenured Faculty

Overview & Reasoning
The task force agreed that the university’s most significant problem with the existing evaluation procedure is how the results are used. Changes to the evaluation instrument and survey process were reviewed (and improvements identified), however no change was found which would produce data sufficiently accurate and valid to be good enough to be the primary or sole measure of teaching effectiveness.

Appraising the problem from the opposite side has the potential to produce better measures of teaching effectiveness, which is the primary way that the university uses existing evaluation data.

This recommendation is intended to improve the way teaching effectiveness is measured by integrating better practices into the university culture.
The above two page document is the full and final report of the Provost’s Task Force on Student Teaching Evaluations. By signing below, I represent that I both worked on generating this report and agree with the report.

Respectfully Submitted:

Lawrence Overzet
Chair

Abhijit Biswas

Gail Breen

Harold Clark

Peter Ingrao

Simon Kâne

Kim Knight

Angela Mcnulty

Sai Nikitha Prattipati
Student Member

Paul Diehl (ex officio)

Karen Huxtable-Jester (ex officio)