

COURSE AND TEACHER SURVEYS (CATS) AT VILLANOVA UNIVERSITY

A guide for new faculty members

There are probably more studies of student ratings than of all of the other data used to evaluate college teaching combined. Although one can find individual studies that support almost any conclusion, for a number of variables there are enough studies to discern trends. In general, student ratings tend to be statistically reliable, valid, and relatively free from bias, probably more so than any other data used for faculty evaluation. Nevertheless, student ratings are only one source of data about teaching and must be used in combination with multiple sources of data if one wishes to make a judgment of all of the components of college teaching. Further, student ratings are data that must be interpreted. We should not confuse a source of data with the evaluators who use the data to make a judgment.¹

Office of Vice President for Academic Affairs
November 14, 2000

¹ Cashin, William E. (1988) "Student Ratings of Teaching: A Summary of the Research," (IDEA Paper no. 20). Manhattan: Kansas State University, Center for Faculty Evaluation and Development.

You will soon be receiving the packets with the materials for administering Villanova's Course And Teacher Survey (CATS). This guide is intended to answer some of the questions that new faculty members have about student surveys. Your chair or program director can also advise you about CATS. If you have additional suggestions or questions about either this guide or the CATS process, please contact John Immerwahr (tel. 610/519-4522).

Why does Villanova survey students about teaching effectiveness?

Generally, we feel that student perceptions are a useful component of understanding teaching effectiveness. The CATS results can give you as a faculty member an additional source of information about how students perceived your classes. They can also give useful information to your chair who can, in turn, make useful suggestions for improving your teaching effectiveness. Ultimately the CATS results are used as one component of the overall evaluation of your teaching, which will be conducted by your colleagues, and your chair. It should be stressed, however, that the CATS results are intended to be a component of this process. Some people inaccurately describe the CATS surveys as "student evaluations." This is incorrect. *At Villanova, only faculty and administrators evaluate teaching effectiveness.* While student perceptions are taken into account in the process of forming those evaluations, the evaluations are made by faculty members and administrators, not by students.

How often and in which classes are the surveys conducted?

The current policy calls for administration of CATS in most courses every semester during the fall and spring semesters. Classes with fewer than three students are not surveyed, and sometimes department chairs will determine that certain classes should not be surveyed. CATS surveys are not done in the summer unless specially requested by chairs or deans.

How are the surveys administered?

When you receive the packet, you will see detailed instructions to you as to how to administer the surveys. Please read them carefully. The general idea is that they should be administered in the last two weeks of the semester, and at the beginning rather than at the end of the class. You should allow the students ten or fifteen minutes (more if there are supplemental questions) to complete the surveys, you should not be in the room when the surveys are being filled out, and you should ask a student to collect them, put them in a sealed envelope and return them to your departmental office. At that point the surveys go directly to the Office of Planning, Training, and Institution Research (OPTIR) where they are processed.

What questions are asked?

CATS is intended to be a minimum survey, and it is assumed that faculty, departments, and colleges may want to add additional questions. A copy of the instrument is included at the end of this guide. As you will see, the instrument has two overall evaluative questions, eleven diagnostic questions, and three demographic questions. It also includes space for additional questions and for open-ended responses.

May I add questions of my own?

There are opportunities for 26 user defined questions on the reverse side of the instrument, lettered from A to Z. You are free to prepare your own additional questions. Some faculty members ask students to comment on the value of individual assignments or texts, or on other specific aspects of their course. These questions could also be open-ended, with students writing comments on a separate sheet. You may, if you wish, send your additional questions to OPTIR (by November 1 for the fall or by March 15 for the spring) and they will put your questions in the packet. If you prefer, you may just give your own questions out on the day you administer the instrument. Tell the students to answer the questions on your sheet in the spaces provided on the back of the CATS form. When you get the results back, you can correlate the tabulated responses with the questions that you asked. Your college or department may also have supplemental questions, so please check this with your chair before you start making up your own questions.

What is the relationship between the grades I give students and the ratings they give me on the CATS forms?

The relationship between student ratings and grades has been extensively studied by social scientists, and there is little evidence for the common view that students trade high student ratings for high grades or punish instructors with high standards. At Villanova, some of the most lenient graders get low scores and some of the most demanding teachers get high scores. On the other hand, students do not respond well to grading practices that they perceive as arbitrary and unfair, and they will definitely express their resentment about what they perceive as unfair grading practices when it comes time to fill out the CATS forms. Rather than trying to give high grades, new faculty members should focus on giving students greater clarity about expectations. One experienced faculty member put it this way: "Students don't expect to get all As and Bs, but they would like to think that courses are arranged so that a student who put the effort into it could get a top grade. They hate courses where it seems impossible to do well, or where they don't really understand what is expected of them."

When will I get the results?

A month or two after the end of the semester, you'll get an e-mail telling you that the CATS results are available on Novasis (the same system where you post grades). The Novasis reports are only available to you, and require the same password system that is used in entering grades. A few weeks after that, the filled out forms themselves (including the results to the open-ended questions) and printed reports will be sent to your chair. The chair will review the forms and then return them to you. Unfortunately, it will be a number of weeks before you get to see the open-ended material.

Who else will see the results?

The CATS reports go to you, your chair, your dean, and to the VPAA. Only you and your chair will see the results of the open-ended questions. New full-time faculty members go through a formal process of evaluation once a year, and teaching is a big part of that evaluation. Typically

chairs have a faculty committee to recommend on the evaluation, and those committees usually have access to CATS reports.

Will the results be available on the world wide web or in the library?

According to a policy passed by the University Senate, the CATS reports may, with the permission of the faculty member, be published either on-line or in the library. When you see the Novasis version of your reports, you are given the option to make the reports available on-line. If you do so, the reports will be available to Villanova students, staff, and faculty who have valid Novasis usernames and PINs. Some students review the published CATS results in deciding which classes to take. The decision to publish your CATS scores is purely up to you, and you may change your decision as often as you want.

You will also receive an e-mail asking if you wish to have your reports made available in the reserve room. If you assent, your reports will be included in a reserve room volume (utilized by few students). This is a one-time decision; once you agree to have your reports in the Reserve Room, the decision is final.

What principles are used in reading the results? Do some classes or subjects get higher ratings than others?

Generally speaking, intro classes, required classes, and classes with a lot of math get lower scores than small, upper division electives. Most of the faculty members, chairs, and administrators who look at the CATS results are rather familiar with these principles and take them into account. At the end of this document we have included the "Principles for use of student ratings for faculty evaluation and development."

After you get your results back, you might want to discuss the results with your chair or program director. The chair can tell you how your CATS surveys will be regarded by your department. Often new faculty members remark that the results they get in the first semester or two are lower than what they saw at their previous institution. After a few semesters, the results frequently improve.

What other information will be considered when my teaching is reviewed?

Your chair can tell you what information is used by your department in evaluating teaching effectiveness. Usually departments look at things such as syllabi, tests, peer observations, and grading patterns. Realistically, chairs and your faculty colleagues also draw on a wide range of information when they come to think about teaching effectiveness. Especially at registration time, chairs spend a great deal of time talking to students about courses, and they hear a lot about student perception of teaching effectiveness. The same is true of other colleagues. The result is that when it comes time to do a formal evaluation of teaching, a lot of factors are taken into account, not just CATS reports.

What can I do to improve my ratings?

New faculty members should concentrate on teaching effectiveness rather than student survey results. The question they should ask themselves is: "What can I do to be a more effective teacher?" Trust us, the ratings will take care of themselves. There are a number of strategies for improving teaching effectiveness.

- ❑ One of the most useful steps is to draw on the resources of the Villanova Institute For Teaching and Learning (VITAL). VITAL offers a host of services ranging from confidential consultations to public workshops.
- ❑ Generally speaking, teaching effectiveness does improve over the first few semesters, as faculty members get a clearer sense of student learning styles.
- ❑ Faculty members may also want to ask advice of recognized good teachers in the department, and may ask permission to observe classes.

Will they be used in determining salary?

In most departments, salary recommendations are based on the annual evaluations, and CATS data are a component in forming the evaluations of teaching.

Will the results be used in decisions regarding rank and tenure?

The CATS reports figure into Rank and Tenure decisions in two ways. First, the Rank and Tenure committees will look at the past departmental annual evaluations. In so far as CATS reports are a component of past teaching evaluations, they influence this aspect of Rank and Tenure. The Rank and Tenure committees also look at a summary report of the applicants CATS scores. Generally, the committee only sees the results for the purely evaluative questions (16 and 17).

Who designed the CATS instrument, can it be changed?

CATS was designed by a committee of faculty and administrators, and was circulated to the entire faculty for comments and revisions. It was inaugurated in 1997. While there have been a number of changes in the administration of CATS, the University has resisted changes to the basic questions themselves, thinking that it was better to stick with the instrument until such time as it was thought to be useful to have a full-scale review.

November 25, 1996

There are probably more studies of student ratings than of all of the other data used to evaluate college teaching combined. Although one can find individual studies that support almost any conclusion, for a number of variables there are enough studies to discern trends. In general, student ratings tend to be statistically reliable, valid, and relatively free from bias, probably more so than any other data used for faculty evaluation. Nevertheless, student ratings are only one source of data about teaching and must be used in combination with multiple sources of data if one wishes to make a judgment of all the components of college teaching. Further, student ratings are data that must be interpreted. We should not confuse a source of data with the evaluators who use the data to make a judgment.

(Cashin, 1988)

Principles for the use of student ratings for faculty evaluation and development

In using and interpreting student ratings, committees, department chairs and other administrators should be guided by these principles:

1. Student ratings are a valuable source of data about effective teaching. (Research shows that they are highly reliable and stable, and are relatively valid against a variety of indicators).
2. Student ratings must be supplemented with other sources of information about teaching, such as: syllabi; tests; grade distributions; faculty self-evaluations; and reports of class observations by colleagues.
3. Student ratings are only one tool for helping to improve teaching. Ratings can be a valuable source of information that aids in instructional improvement. They may help to identify particular aspects of a course that merit attention (e.g., test and exams, organization, student interaction, etc.). However, while serving this "diagnostic" function, standard rating forms are not very "prescriptive" in terms of specific suggestions to the instructor for making adjustments. In this regard, the role of one's peers or chairperson, confidential interviews with students, responses to open-ended questions, and teaching workshops are probably more valuable. Faculty members should be encouraged to do their own confidential surveys during the course of the semester.
4. Administrators and peer committees who rely on student ratings have a responsibility to be educated in the interpretation of these data. Extensive research has been done on student ratings and those who use these data should be aware of the strengths and limitations of student ratings. Among other things, the research suggests the following principles:
 - ❑ Standardized procedures should be carefully developed and scrupulously followed for the administration of student rating forms. For example, the forms should be anonymous, the instructor should not be in the room when the forms are administered; the forms should be distributed toward the end of the semester; and students should not be unduly rushed when completing them.

- ❑ When using student ratings for evaluative purposes, look primarily at data that include multiple classes or sections across several semesters, and from classes that have at least 10 raters and two-thirds of the class present. Data on a few classes or a few students are likely to be unreliable.
- ❑ Do not overanalyze the data. Small differences between one instructor or another or between one section and another may be meaningless.
- ❑ Look at the results in terms of relevant comparisons; the absolute numbers in student ratings are of only limited value. Department chairs should furnish faculty members with reports that allow faculty members to compare their own results to that of the department as a whole.
- ❑ Take into account the fact that some courses typically receive higher or lower evaluations, and do not inappropriately reward or penalize instructors for these differences. For example, larger classes typically receive slightly lower ratings, as do required courses and courses outside of a student's major area. Upper-level courses (especially graduate level courses) tend to receive higher ratings.
- ❑ When using student ratings for evaluative purposes, focus primarily on "global" evaluative items about overall instructor performance and overall course satisfaction. While these questions are appropriate for evaluative purposes, they are of much less value for developmental purposes. Chairs and others who work with faculty members to help them improve their teaching should rely primarily on the more diagnostic questions.²
- ❑ Allow faculty to submit comments on their student evaluations, and take into account special factors (e.g., whether the course is a new or redesigned course or is team-taught).

Recommended background reading:

Cashin, William E. (1988) "Student Ratings of Teaching: A summary of the Research," (IDEA Paper No. 20). Manhattan: Kansas State University, Center for Faculty Evaluation and Development, and William E. Cashin (1990) "Student Ratings of Teaching: Recommendations for Use," (IDEA Paper No. 22). Manhattan: Kansas State University, Center for Faculty Evaluation and Development.

VPAA November 25, 1996

² This finding is somewhat counterintuitive; many faculty members criticize evaluation committees for focusing on only one or two questions. The research suggests, however, that if the concern is evaluation, the most global questions are the most reliable indicators.