

**Meeting of the Villanova University  
Academic Policy Committee**

Monday, March 19, 2018  
11:30 - 12:30 PM  
Fedigan Room (SAC 400)

Present: Danai Chasaki, Marylu Hill, Christopher Kilby (chair), Adele Lindenmeyr , Eric Lomazoff, Peggy Lyons, Wen Mao, Emily McCloskey, Elizabeth Petit de Mange, Michael Posner, Rees Rankin, Joseph Schick, Marguerite Schlag, Andrea Welker, Craig Wheeland

Absent: Sherry Bowen (on leave), Sepideh Cheheltani (NIA), Gordon Coonfield (NIA), Jennifer Dixon (NIA), Shelly Howton (on leave), Brian King (NIA), Kimberly Marucci, Dennis Wykoff (on leave), Tina Yang (NIA)

Administrative Items

- 1) Minutes from 2/21/2018 approved with one abstention.

Old Business

- 2) Online CATS Subcommittee (Michael Posner, Chair)

Michael Posner called people's attention to the draft of a revised charge for the subcommittee (including a name changed dropping "Online") and the draft faculty survey that he had emailed to APC (and attached to the agenda). Michael suggested that these items be reviewed at the next APC meeting but that this meeting focus on analysis of CATS and pilot results.

Trina Das (Director for Policy and Planning Analysis, OPIR) presented a PowerPoint summary of analysis of recent CATS, focusing on a comparison of the first full online CATS in Spring 2017 with the paper CATS in Spring 2016. Trina's presentation covered response rates and average scores. As previously documented, overall response rates dropped by about 7%. This drop was similar when broken down by instructor gender and tenure status; the drop is somewhat larger for part time faculty as compared to full time faculty (though this may reflect differences in the length of the classes the two groups teach; see next item). In the Spring 2017 data, there is a strong link between length of class (measured by hours of instruction) and response rates, with full semester length classes having notably higher response rates. (Data on course length are not readily obtainable for paper CATS so no comparison was possible.) Christopher Kilby requested that this analysis be repeated by date of CATS window to see if low response rate is due to length-of-class or time-of-semester effects, since outreach to address low response rates might differ dependent on the true cause.

Turning to rating averages from the Likert scale questions, Trina focused on five questions used in rank and tenure decisions (and typically in other evaluations): Q22, Q23, Q26, Q28, and Q29. Going from paper CATS in Spring 2016 to Online CATS in Spring 2017, few of the differences were statistically significant and even fewer were substantively significant (e.g., greater than 0.1 on the 1 to 5 scale). Craig Wheeland (Vice Provost for Academics) pointed out that deviations less than  $\pm 0.2$  from department or college means are not treated as meaningful in University guidelines for evaluation. There were small differences for Question 22 (lower in Spring 2017) and Question 29 (higher in Spring 2017). When examined by college, the Q22 drop was more notable in CON; OPIR plans to explore this more carefully. Differences from Spring 2016 to Spring 2017 broken down by gender, tenure status, and full time/part time groupings did not reveal notable findings.

Christopher asked about data on what devices students used to complete CATS online (with an eye toward how this impacts survey completion, ratings, and the number and length of comments). If problems are identified, certain types of devices could be prohibited. Trina indicated OPIR would follow-up with the vendor to see if they could provide the device type for each survey.

Ken Tsang (OPIR) presented a PowerPoint summary of analysis of the Fall 2017 pilot of diversity and inclusion CATS questions. There were two Likert scale questions (DI\_Q1: demonstrates cultural awareness; DI\_Q2: bias-free environment) and one opened comment that prompted students with a range of diversity categories. Of the roughly 2200 courses (identified by course registration numbers, CRNs), 157 CRNs were randomly selected; this included 126 different instructors (some instructors included in more than one CRN covered). The 157 CRNs resulted in about 2800 possible CATS (i.e., average class size of about 18); this included 2200 students (some students included in more than one CRN covered). Approximately half of the surveys included an "NA" option with the Likert scale questions. Students only rarely selected this option (3.7% for DI\_Q1 and 2% for DI\_Q2) and rating averages for the groups with and without the "NA" option did not differ. Comparing responses on non-D&I questions for CATS with and without D&I questions, there were no apparent differences (i.e., inclusion of pilot questions did not appear to impact results for Q22, Q23, Q26, Q28, and Q29 and so had no adverse impact on faculty selected for the pilot).

There were some differences in rating averages by college, with somewhat lower scores for VSB and Sciences than for ENG, Humanities/Social Sciences, and CON. However other than for Economics, Ken reported that there was no college variation in students taking each course. That is, Humanities/Social Sciences courses included only Humanities/Social Science students (no Science students, etc.), VSB courses included only VSB students, etc. Several APC members thought this unlikely; Trina and Ken indicated that OPIR would recheck how the data were coded to see if this was an error.

Taking a cue from the existing literature, Ken broke the data down by instructor and student gender. For DI\_Q2, scores were highest for female student-female instructor pairs (4.63), very slightly lower (and not statistically significantly so) for male-female and male-male pairs, and lowest for female student-male instructor pairs (4.49). However, from current data it is not possible to determine the cause of this difference (i.e., instructor bias or student bias). This pattern was evident only for DI\_Q2.

Alternately breaking the data down by race and ethnicity, there were not clear differences. It was not possible to break this down further (e.g., within colleges) due to small sample size.

As for the open-ended question, about 1/3 of respondents entered comments. Responses were generally consistent with the numerical scores. Some students commented that the questions were not relevant given the course material; this was more common for STEM courses. There was some clustering of comments/lower scores by faculty member, suggesting that is useful information in these student responses.

Michael raised the question of how survey responses might be used, since this may be an important consideration in the design of the questions. Michael suggested that chairs might discuss diversity and inclusion issues with faculty if the number of negative comments/responses passes some threshold (e.g., 5). Christopher suggested the percent negative comments/responses from minority students—however OPIR is able to capture that—might be another consideration. [On reflection, this is probably not possible while preserving confidentiality.] Marguerite Schlag suggested that even one negative comment/rating is reason enough for a chair to discuss the issue with the faculty member.

Terry Nance strongly emphasized the need to move forward with the D&I CATS questions so that the university has internal data to assess where it stands more concretely than is possible now. Wen Mao pointed out the utility of such data went addressing specific complaints, as it would allow administrators to see whether there is corroborating evidence from CATS.

Christopher emphasized that, going forward, OPIR—through its vendor—needs to have the ability to do CATS pilots as separate modules (after responses to regular questions have been recorded and cannot be changed) and that faculty have the right to see questions (pilot or otherwise) before their students answer these questions and to see the results (even if these results are not used or seen by others).

The meeting adjourned at 12:43 PM.