

UNIVERSITY OF OTAGO INTEROFFICE MEMORANDUM

FROM Allen Goodchild, Evaluation Coordinator, Quality Advancement Unit  
 SUBJECT Additional metrics in student evaluation of teaching summary reports  
 DATE 16 January 2020

The revised student evaluation of teaching summary report retains all of the information from previous years' reports (yellow in below image). The revisions add metrics for question 5 (Overall, how effective have you found Dr Spock in teaching this course?) and response rates (blue in below image).

<b>Dr Jane Spock</b> Evaluations of teaching: summary data University of Otago <b>Generic Questions (compulsory)</b>	School of Physical Education, Sport and Exercise Sciences  2019 Teaching Type Ratings (1 & 2 Preferred)	<b>Test - checking reports (Group 1)</b> Date: 20-22 Jun 2019 Responses: 11 % Class: 44 Distance 1&2 3 4&5	<b>Test - checking reports (Group 2)</b> Date: 20-22 Jun 2019 Responses: 6 % Class: 60 Distance 1&2 3 4&5	Date Res % C 1&
		100 0 0 100 0 0 100 0 0 100 0 0 55 36 9	100 0 0 50 50 0 17 0 83 100 0 0 83 17 0	100 0 0 50 50 0 17 0 83 100 0 0 83 17 0

- (1) Q5 in more detail: 'Overall, how effective have you found Dr Spock in teaching this course?'
 

95% confidence interval (CI) for proportion of 1s and 2s:	32% to 72%	60% to 90%
---	------------	------------
- (2) Response rates
 

Percentage of comparison classes with the same or lower response rates:	56%	55%
---	-----	-----
- (3) Comparison classes
 

Similar sized classes between 2015 and 2018: [C = Commerce, HS = Health Sciences, H = Humanities, S = Sciences, U = University]	1587 S classes, 21 to 50 students	2274 S classes, 0 to 20 students
--	--------------------------------------	-------------------------------------
- (4) Combined ratings for Q5 during 2019
 

95% CI for proportion of 1s and 2s combining 17 ratings from 2 evaluations:	52% to 79%	
---	------------	--

(1) Q5 in more detail: 'Overall, how effective have you found Dr Spock in teaching this course?'  
 This 95% confidence interval (CI) is a way of representing the entire class when some students don't respond. It means that if the questionnaire was re-run with 100 classes, we could expect 95 of the CIs to contain the actual proportion of positive teacher effectiveness ratings (1s and 2s) if the entire class had responded<sup>1</sup>. As the number who respond increases, the CI narrows to reflect greater precision. And, if the entire class responds, there is no CI because the percentage of 1 and 2

<sup>1</sup> Though we call it a confidence interval, the Bayesian approach used to calculate it calls it a credible interval. This CI is calculated using the finite population correction factor.

responses to question 5 (q5) would already represent the entire class. The average CI width for q5 results between 2017 and 2019 was 28%.

This CI estimation assumes that both questionnaire responders and non-responders hold similar views, which we are unsure about. But our University data shows that, in general, students who had higher grades in a paper were more likely to respond to its questionnaire, and secondly, that those with higher grades rated teachers as more effective than those with lower grades. So, it may be that the CI will more positively portray how the entire class would have responded.

## (2) Response rates

This is the percentage of comparison classes (as defined below) that obtained either the same or a lower response rate than the response rates obtained in the questionnaires reported here.

## (3) Comparison classes

The classes used as a comparison are from the 3 calendar years prior to the start year of the questionnaire reported on and are filtered on the division of the staff member (according to our records) and class size.

Limiting the comparison classes to the previous 3 calendar years is somewhat arbitrary, but it matches both the period of a standard degree, and the period for which HR's Academic Staff Promotions policy requires applicants to provide all their teacher evaluation reports from. It also allows for the comparison group to evolve by a third each year and thus reflect trends more quickly than a longer period. A shorter period would sharply reduce the number of classes in each comparison group once it is also filtered by division and class size.

The comparison group is filtered by division to allow a sufficiently large pool of somewhat similar classes, and by class size because response rates tend to decrease as class size increases.

## (4) Combined ratings for Q5 during 2019

This 95% CI is based on combining q5 responses to all individual evaluation reports across the year. By aggregating responses, it aims to provide a narrower/more precise CI than multiple CIs for individual evaluations in which only a few students respond. With each report given an equal weighting, this combined CI may be good for representing multiple evaluations of similar classes/streams. However, where quite different teaching situations are reported on, the combined CI may hide a more complex picture which is better told by reflecting upon just the individual reports.

For more detail about any of these changes to the teacher evaluation summary report, please contact either Julie or Allen at the Quality Advancement Unit, [evaluation@otago.ac.nz](mailto:evaluation@otago.ac.nz) or (03) 479 7581.