

# An Analysis of Winter 2024 Student Course Perceptions Survey Responses



ASSOCIATE VICE-PRESIDENT, ACADEMIC | Teaching Assessment  
Processes [uwaterloo.ca/teaching-assessment-processes](https://uwaterloo.ca/teaching-assessment-processes)

## **An Analysis of Winter 2024 Student Course Perceptions Survey Responses**

Prepared January 2026 by:

Sonya Buffone, Director, Teaching Assessment Processes

Kathy Becker, Specialist, Teaching Assessment Processes

Joslin Goh, Senior Statistician, Statistical Consulting and Survey Research Unit

# An Analysis of Winter 2024 Student Course Perceptions Survey Responses

## Contents

|                                                                                                      |    |
|------------------------------------------------------------------------------------------------------|----|
| Executive Summary.....                                                                               | 4  |
| Introduction.....                                                                                    | 6  |
| Sources of Data.....                                                                                 | 6  |
| Research Aims .....                                                                                  | 6  |
| <b>Primary Analyses</b> .....                                                                        | 7  |
| Key Findings.....                                                                                    | 7  |
| Descriptive Analysis of Demographics and Course Characteristics.....                                 | 7  |
| Instructor-level Variable: Indigenous Identity.....                                                  | 7  |
| Instructor-level Variables: Racial Identity and Time in Canada.....                                  | 8  |
| Instructor Racial Identity .....                                                                     | 8  |
| Racialized Instructors by Faculty.....                                                               | 9  |
| Instructor Time in Canada.....                                                                       | 10 |
| Instructor Racial Identity by Time in Canada .....                                                   | 11 |
| Instructor-level Variable: Sex.....                                                                  | 12 |
| Instructor Sex .....                                                                                 | 12 |
| Instructor Sex by Faculty .....                                                                      | 13 |
| Instructor Sex by Appointment Type .....                                                             | 14 |
| Course-level Variable: Class Size .....                                                              | 15 |
| Course-level Variable: Course Type by Faculty.....                                                   | 16 |
| SCP Responses by Response Item .....                                                                 | 17 |
| SCP Mean Ratings .....                                                                               | 18 |
| Mean Ratings by Response Item and Class Size .....                                                   | 19 |
| Mean Ratings by Response Item and Faculty of Course Offering.....                                    | 20 |
| Differences in Mean Ratings .....                                                                    | 21 |
| Analysis of Differences: Instructor Sex.....                                                         | 22 |
| Difference in MR for Male and Female Instructors.....                                                | 22 |
| Difference in MR Assigned to Male and Female Instructors by Faculty of Course Offering .....         | 24 |
| Difference in MR for Male and Female Instructors by Faculty of Course Offering and Class Size .....  | 25 |
| Difference in MR for Male and Female Instructors by Faculty of Course Offering and Course Type.....  | 26 |
| Difference in MR for Male and Female Instructors by Class Size and Instructor Appointment Type ..... | 27 |
| Difference in MR for Male and Female Instructors by Class Size .....                                 | 28 |
| Difference in MR for Male and Female Instructors by Course Type, and Instructor Appointment Type     | 29 |

## An Analysis of Winter 2024 Student Course Perceptions Survey Responses

|                                                                                                                 |    |
|-----------------------------------------------------------------------------------------------------------------|----|
| Difference in MR for Male and Female Instructors by Faculty of Course Offering and Instructor Appointment ..... | 30 |
| Difference in MR for Male and Female Instructors by Course Type and Class Size .....                            | 31 |
| Recap of Key Findings.....                                                                                      | 32 |
| Concluding Remarks.....                                                                                         | 34 |
| Moving Forward: Implications for Teaching Assessment Practice .....                                             | 35 |

## Executive Summary

This report presents the University of Waterloo's third institutional analysis of Student Course Perceptions (SCP) survey data following the launch of the revised SCP instrument in Winter 2022. Drawing on 47,737 student responses across 2,440 section-instructor pairs from Winter 2024, the analysis examines patterns in SCP ratings in relation to selected instructor-level and course-level characteristics. The study forms part of the university's broader shift toward a holistic, evidence-informed approach to teaching assessment.

Winter 2024 SCP mean ratings demonstrate a high degree of consistency with those observed in Winter 2022 and Winter 2023. Across all six core SCP items, mean ratings remain strongly skewed toward the high end of the five-point scale, with the majority of responses falling in the "Agree" and "Strongly Agree" categories. The SCP continues to function as a stable and interpretable instrument for capturing broad patterns in student perceptions of their learning experiences.

Class size remains the most consistent correlate of SCP ratings, with students rating smaller classes slightly higher across all items than larger classes. These differences are modest in magnitude and align with existing literature, prior Waterloo SCP analyses, and the common finding that students in smaller classes report a better learning experience than those in larger classes. Differences across Faculties are similarly small, with mean ratings clustering tightly within a narrow range. Across Faculties and class sizes, items related to learning environment, clarity of concepts, and learning outcomes tend to receive higher ratings than items related to stimulating interest and course activities.

Instructor-level analyses were constrained by substantial nonresponse to the institutional Equity Survey. As a result, analyses examining SCP ratings by Indigenous identity, racial identity, and gender identity could not be conducted for Winter 2024 without risking unreliable inferences or compromised anonymity. Available equity data are therefore presented descriptively only.

To allow comparative analyses despite the lack of gender identity data, sex at birth was used instead. Exploratory analyses of differences in mean ratings assigned to male and female instructors indicate that sex-based differences in SCP ratings are generally small and overwhelmingly not statistically significant. Where statistically significant differences do appear, typically within specific combinations of Faculty, class size, or appointment type, the magnitude of these differences remains modest and well under half a point on the five-point Likert scale. These findings are consistent with previous SCP analyses and reinforce the need for cautious interpretation, particularly in the context of multiple comparisons and small subgroup sizes.

Taken together, the Winter 2024 results underscore the stability of the SCP survey over time and highlight the limited interpretive value of small numerical differences when considered in isolation. The findings reinforce the importance of using SCP data as one component within a broader, multi-measure framework for teaching assessment. Within such a framework, SCP ratings can help identify broad patterns, while open-ended student comments, peer review of

## An Analysis of Winter 2024 Student Course Perceptions Survey Responses

teaching, and teaching dossiers provide the necessary context to support fair, meaningful, and development-oriented assessment of teaching effectiveness.

## Introduction

As part of a broader operational and cultural shift to a holistic model of teaching assessment at the University of Waterloo, the Student Course Perceptions (SCP) Survey was officially launched in Winter 2022. The SCP survey is the first step in implementing this holistic model, which is grounded in institutional teaching and learning priorities. As part of its commitment to the ongoing monitoring of the SCP survey, the Teaching Assessment Processes (TAP) office collaborated with the Statistical Consulting and Survey Research Unit (SCSRU) in the Faculty of Mathematics to analyze SCP data from Winter 2024. The present study is the third in the series, with the first being An Analysis of Winter 2022 SCP Survey Responses. All previous reports can be accessed here: <https://uwaterloo.ca/associate-vice-president-academic/reports-uw-community>

## Sources of Data

Data for this analysis was drawn from three primary sources:

1. Winter 2024 SCP survey data and, for instructors whose Winter 2024 course(s) had SCP surveys administered,
2. Waterloo's Equity Census data, and;
3. Human Resources records on 'sex at birth.'

An important limitation to this study relates to the high frequency of missing data for racial identity; in total, 799 of the 1,398 instructors in this study did not provide information for the racial identity question on the Equity Survey. This represents 57.2% of instructors with missing equity data for this question, an increase from 45.7% in the Winter 2023 dataset. Two main factors likely contributed to this decrease in response rates: the transition of the Equity Survey from one platform to another (from Qualtrics to the Workday platform) and reported concerns among instructors about providing sensitive personal demographic information through an HR platform, as systems designed for employment administration may not feel appropriate for voluntary self-identification.

Due to the high proportion of missing responses for this survey question, the available equity data for Winter 2024 did not meet the threshold for reliable or representative analysis. Proceeding with subgroup comparisons under these conditions would risk both compromising anonymity for smaller groups and drawing conclusions that are not methodologically sound. For these reasons, analyses related to instructor Indigenous identity, racial identity, and gender identity were not conducted for this dataset; demographic data are included for descriptive purposes only.

## Research Aims

The analysis in this report was guided by the following research focus: to identify the extent to which key factors identified in the literature, and by our Waterloo community members, are correlated with SCP ratings assigned by student respondents. Specifically, this analysis explored the strength of associations between SCP ratings and:

## An Analysis of Winter 2024 Student Course Perceptions Survey Responses

1. Instructor-Level Variables: instructor sex, instructor appointment, and;
2. Course-Level Variables: class size, course type (i.e., online or in-person) and Faculty of course offering.

A detailed coding scheme is provided in Appendix A.

### Primary Analyses

The report consists of three primary analyses:

1. A descriptive analysis of instructor demographics and course characteristics at Waterloo for Winter 2024.
2. An analysis of the mean SCP ratings across different combinations of instructor-level and course-level variables.
3. An analysis of the difference in mean ratings between male and female instructors across different combinations of instructor-level and course-level variables.

### Key Findings

Due to the above-noted nonresponse (missing cases) in equity data, we were unable to examine the impact of Indigenous identity or racial identity on SCP ratings since data on instructor Indigenous identity was both incomplete and small.

## Descriptive Analysis of Demographics and Course Characteristics

In total, 47,737 SCP surveys from 2440 section-instructor pairs were submitted by students from across the six Faculties and Renison. It is important to emphasize that, since one section of a course might have more than one instructor, and since one instructor might teach more than one section, the 2459 section-instructor pairs include the following:

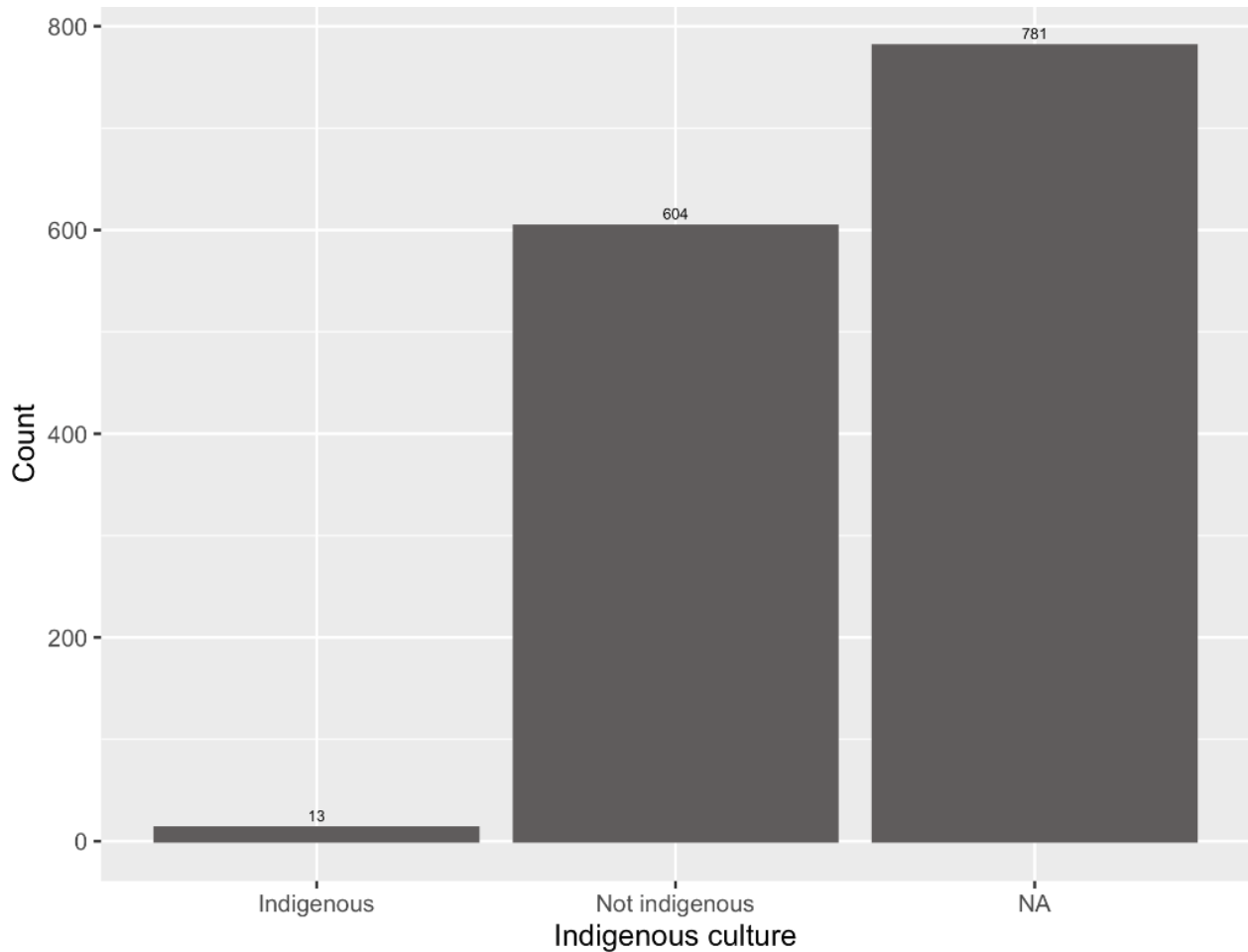
- 1398 unique instructors
- 1601 unique courses
- 2321 distinct sections

Due to issues the team faced when merging various data sources, some instructor-section pairs may be counted more than once in cases where they are cross-listed in different faculties and have different course-level information listed. Such cases were merged in the data set as a result.

### Instructor-level Variable: Indigenous Identity

Thirteen Winter 2024 instructors identified as Indigenous in the Equity Survey (see Figure 1).

## An Analysis of Winter 2024 Student Course Perceptions Survey Responses



*Figure 1: Number of Winter 2024 instructors by self-reported Indigenous identity – Note: Each Equity Survey respondent was asked to select all applicable racial identities. NA represents Winter 2024 instructors who did not complete the Equity Survey or preferred not to answer this specific question.*

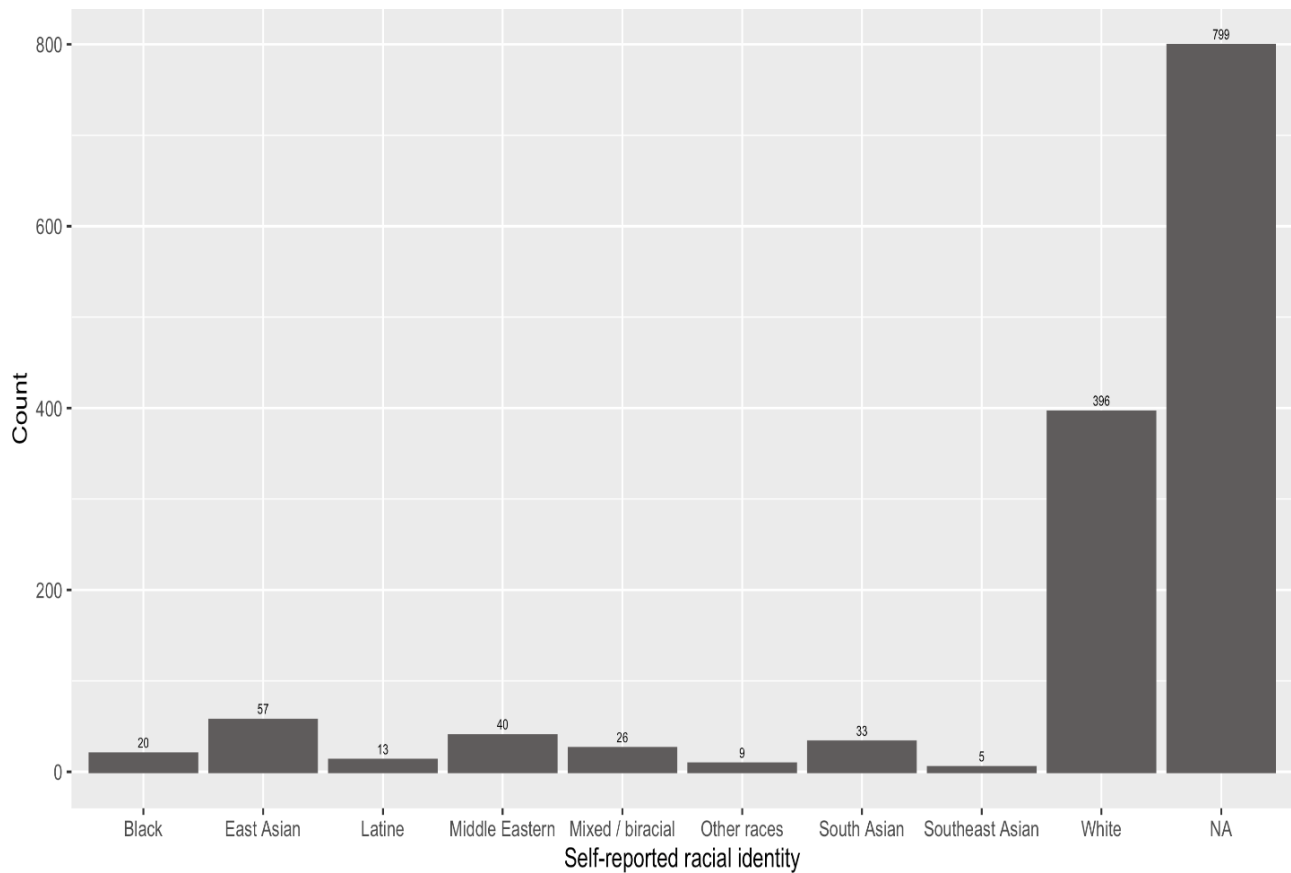
## Instructor-level Variables: Racial Identity and Time in Canada

### Instructor Racial Identity

Turning to Figure 2, a further breakdown of the equity data shows the racial identities of Winter 2024 instructors. In cases where a respondent selected more than one racial identity, they were coded in each of the categories selected. For example, an instructor who selected both “East Asian” and “Another race category” would be included in those two racial identity categories.

As Figure 2 illustrates, there is a significant amount of missing data for this variable with 799 instructors not providing any racial data. Based on the responses we received, the single largest racial category for Winter 2024 instructors was white (396), followed by 57 East Asian, 40 Middle Eastern, and 33 instructors indicating they are South Asian. Only 20 instructors who provided racial data indicated they are black.

## An Analysis of Winter 2024 Student Course Perceptions Survey Responses



*Figure 2: Number of Winter 2024 instructors by self-reported racial identity - Note: Each Equity Survey respondent was asked to select all applicable racial identities. NA represents Winter 2024 instructors who did not complete the Equity Survey or preferred not to answer the racial identity question.*

### Racialized Instructors by Faculty

We examined the number of racialized instructors across each Faculty who had SCP surveys in Winter 2024. As shown in Figure 3, there is a high number of instructors who did not report racial data across all six Faculties. Setting missing cases aside, there were nearly three times more white than racialized instructors teaching in Winter 2024 across three of the six Faculties (Arts, Environment, Math). Just under two times more white than racialized instructors in one Faculty (Science), and almost equal numbers of white and racialized instructors in one Faculty (Engineering).

## An Analysis of Winter 2024 Student Course Perceptions Survey Responses

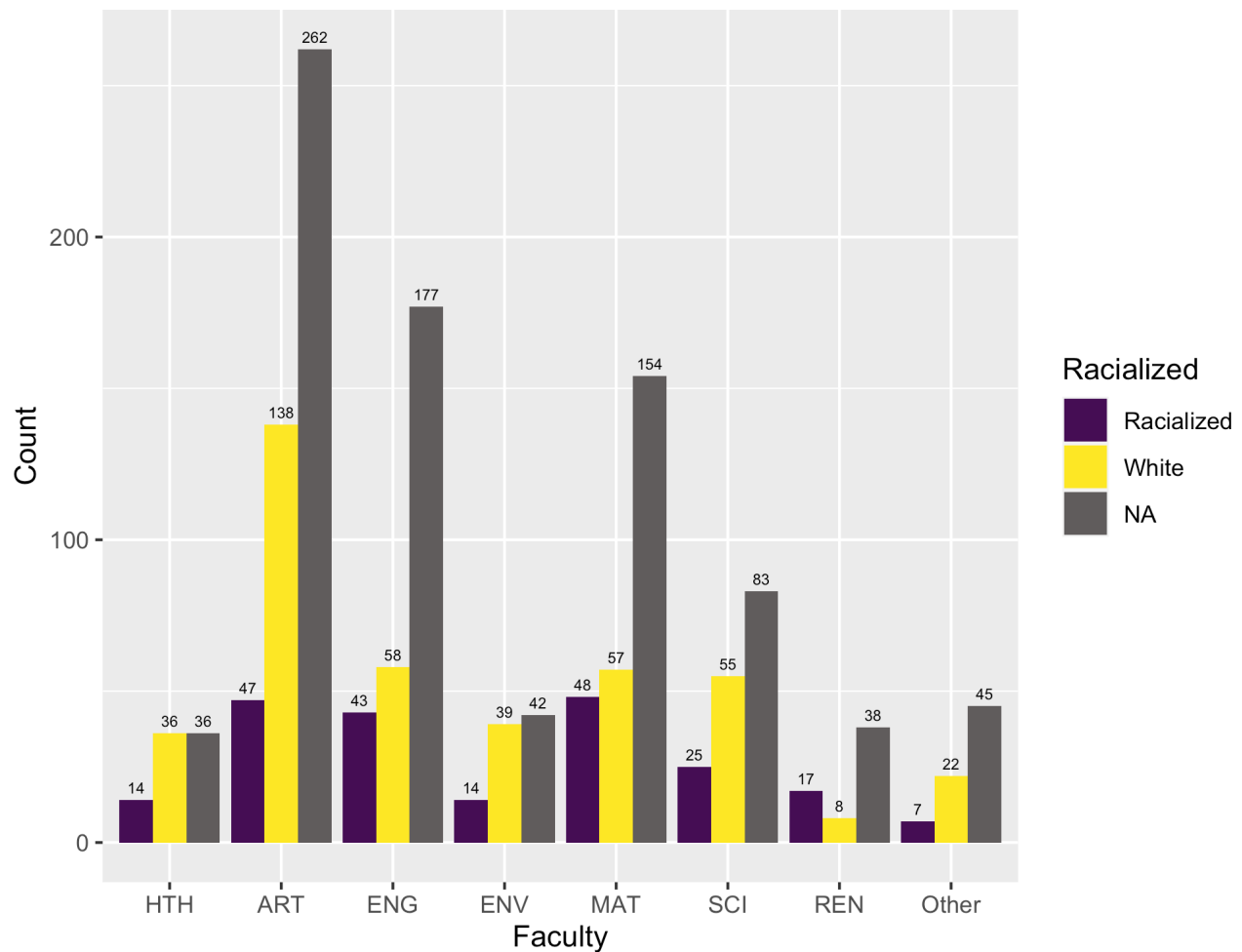


Figure 3: Number of Winter 2024 instructors by faculty and racial identity – Note: NA represents Winter 2024 instructors who did not complete the Equity Survey or preferred not to answer the racial identity question. “Other” includes University Colleges (except for Renison) and those that do not belong to any Faculty (e.g., GRAD).

### Instructor Time in Canada

Figure 4 outlines the length of time in Canada for Winter 2024 instructors who responded to the Equity Survey. Overall, 338 instructors in the sample were born in Canada, 252 had been in

## An Analysis of Winter 2024 Student Course Perceptions Survey Responses

Canada 5+ years and only 44 instructors reported being in Canada for less than 5 years. Again, we highlight there are 764 instructors who did not provide any data for this question.

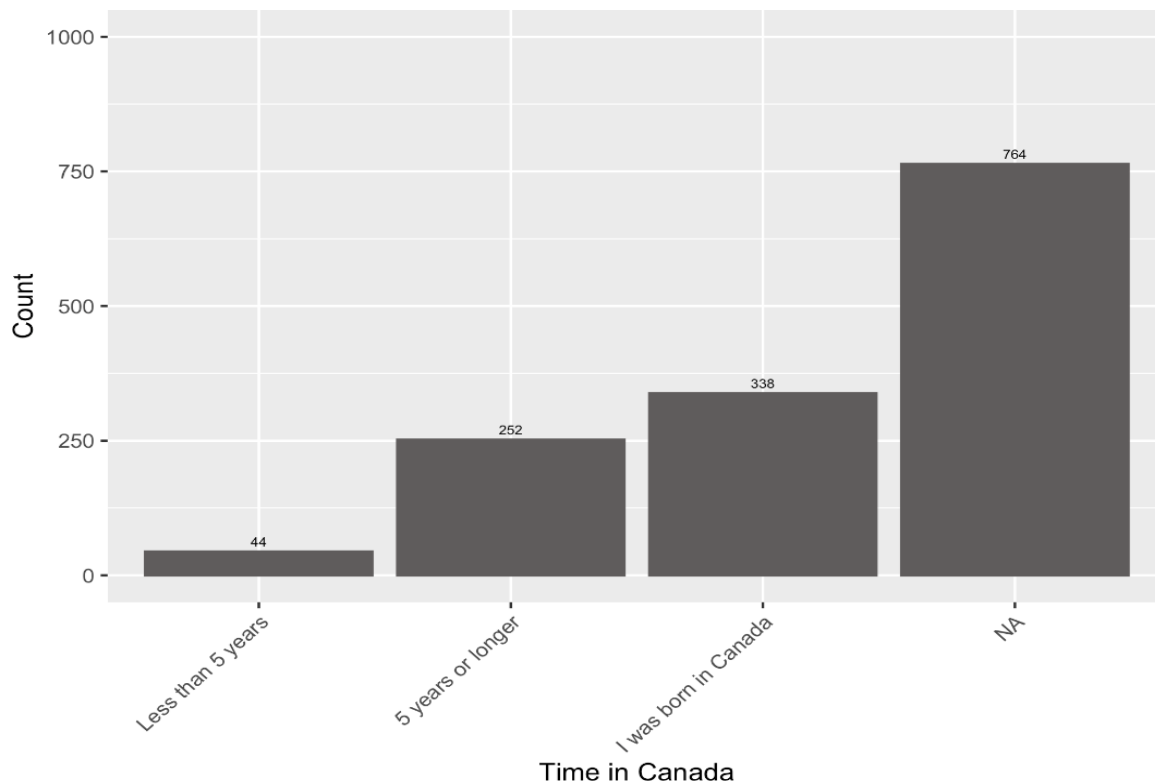


Figure 4: Number of Winter 2024 instructors by self-reported time in Canada. Note: NA represents Winter 2024 instructors who did not complete the Equity Survey or preferred not to answer the time in Canada question.

### Instructor Racial Identity by Time in Canada

Figure 5 provides a further breakdown of the characteristics of instructors who responded to the Equity survey. As indicated by the yellow bars, we can see that of the respondents who are white, over two-thirds (71%) were born in Canada. This is in contrast to a smaller number of racialized respondents (19%) who were born in Canada. For those respondents who were born outside of Canada, two-thirds of those who are also racialized (67%) report being in Canada 5+ years compared to only one quarter of those who are white (25%) report being in Canada 5+ years.

## An Analysis of Winter 2024 Student Course Perceptions Survey Responses

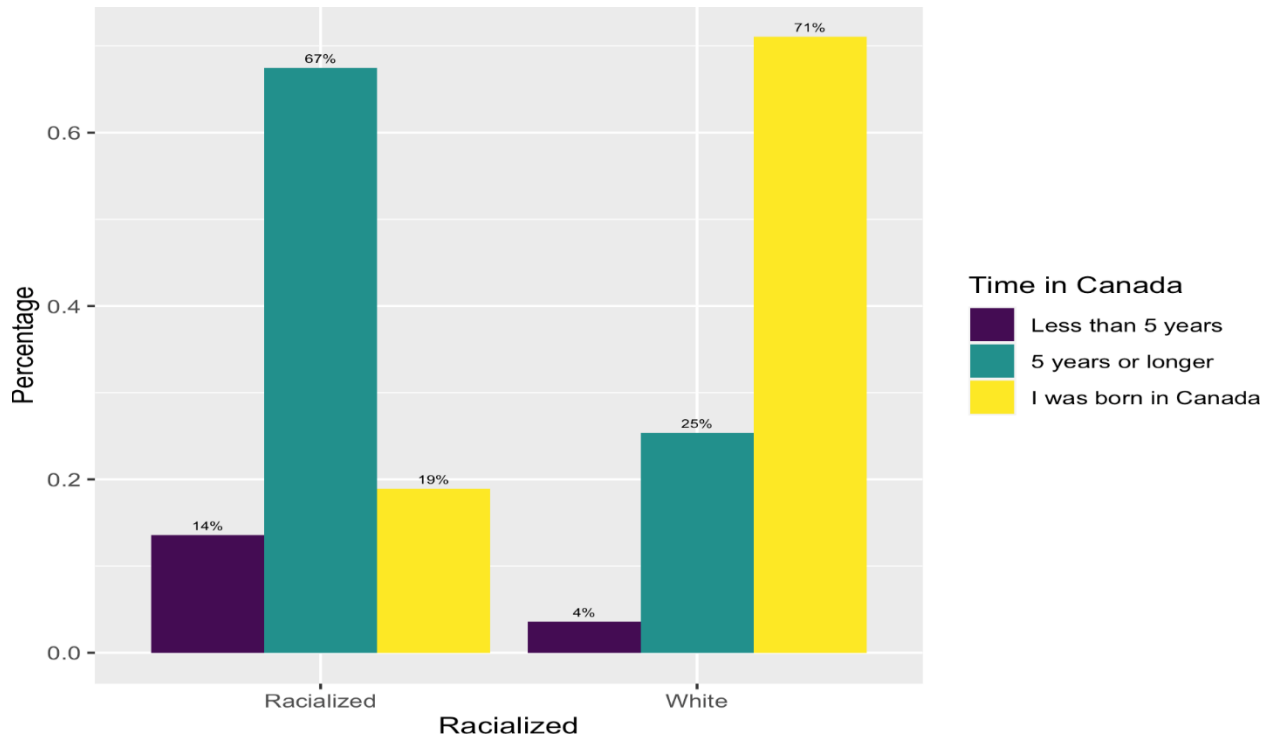


Figure 5: Self-reported Number of white and racialized Winter 2024 instructors by time spent in Canada.

## Instructor-level Variable: Sex

### Instructor Sex

Figure 6 displays Human Resources data for 'sex at birth' for Winter 2024 instructors with 39.4% of the sample female and the remaining 60.5% of the sample male (551 female and 847 male). It's important to note that this data does not necessarily reflect instructor gender identity at the present time.

## An Analysis of Winter 2024 Student Course Perceptions Survey Responses

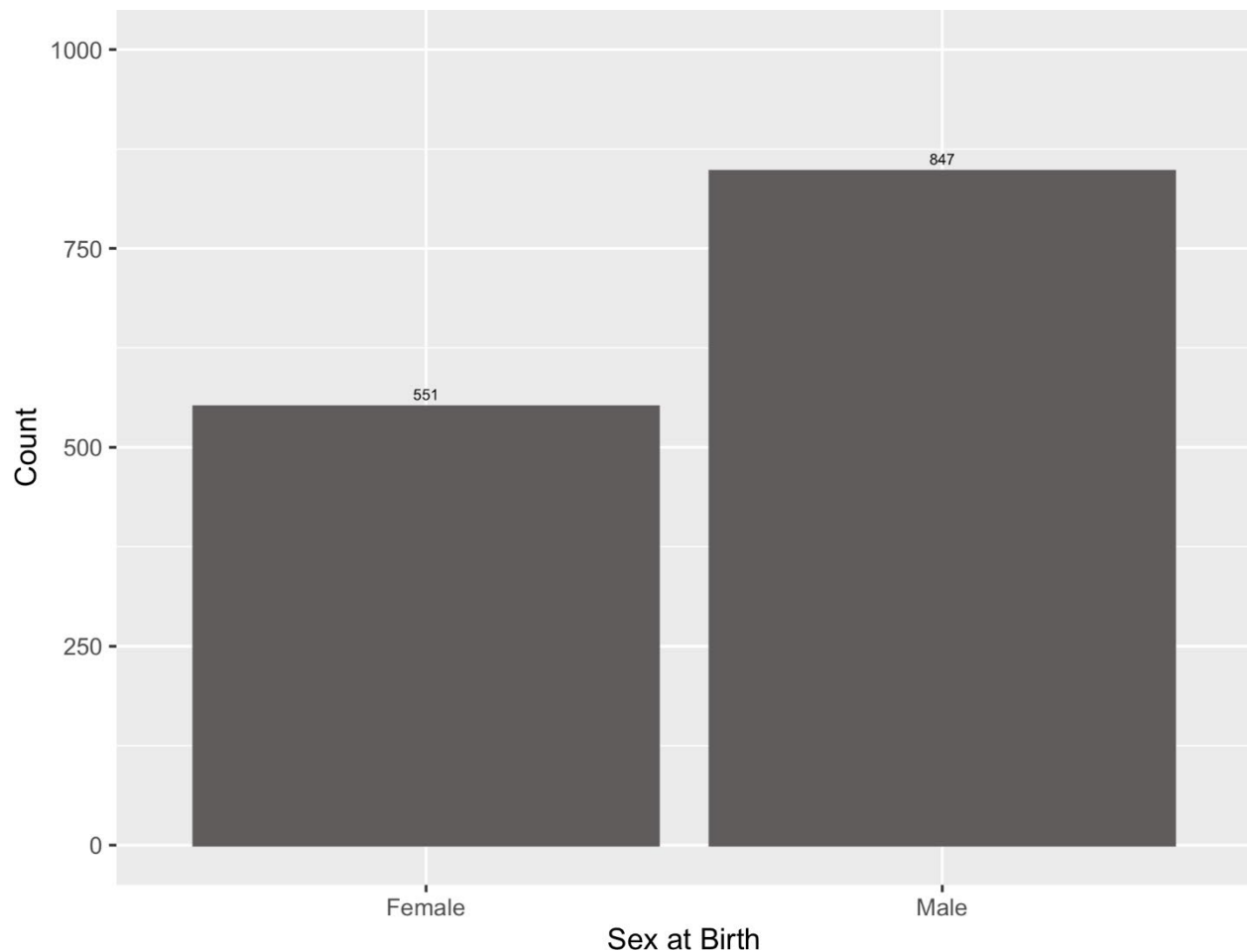


Figure 6: 2024 Human Resources data for Sex at Birth

### Instructor Sex by Faculty

Figure 7 shows the sex at birth for Winter 2024 instructors across each faculty. Male instructors outnumber female instructors across the STEM Faculties: Engineering, Math and Science. In two STEM Faculties, there are over three times more male than female instructors (Engineering: 211 males, 67 females, and Math: 195 males, 64 females). In Science, we see just under two times more male than female instructors (95 males and 68 females). Similarly, in Environment, we see just under two times more male than female instructors (Environment: 56 males, 39 females). And in one Faculty we see near parity in the number of male and female instructors teaching in Winter 2024 (Arts: 238 males, 209 females).

## An Analysis of Winter 2024 Student Course Perceptions Survey Responses

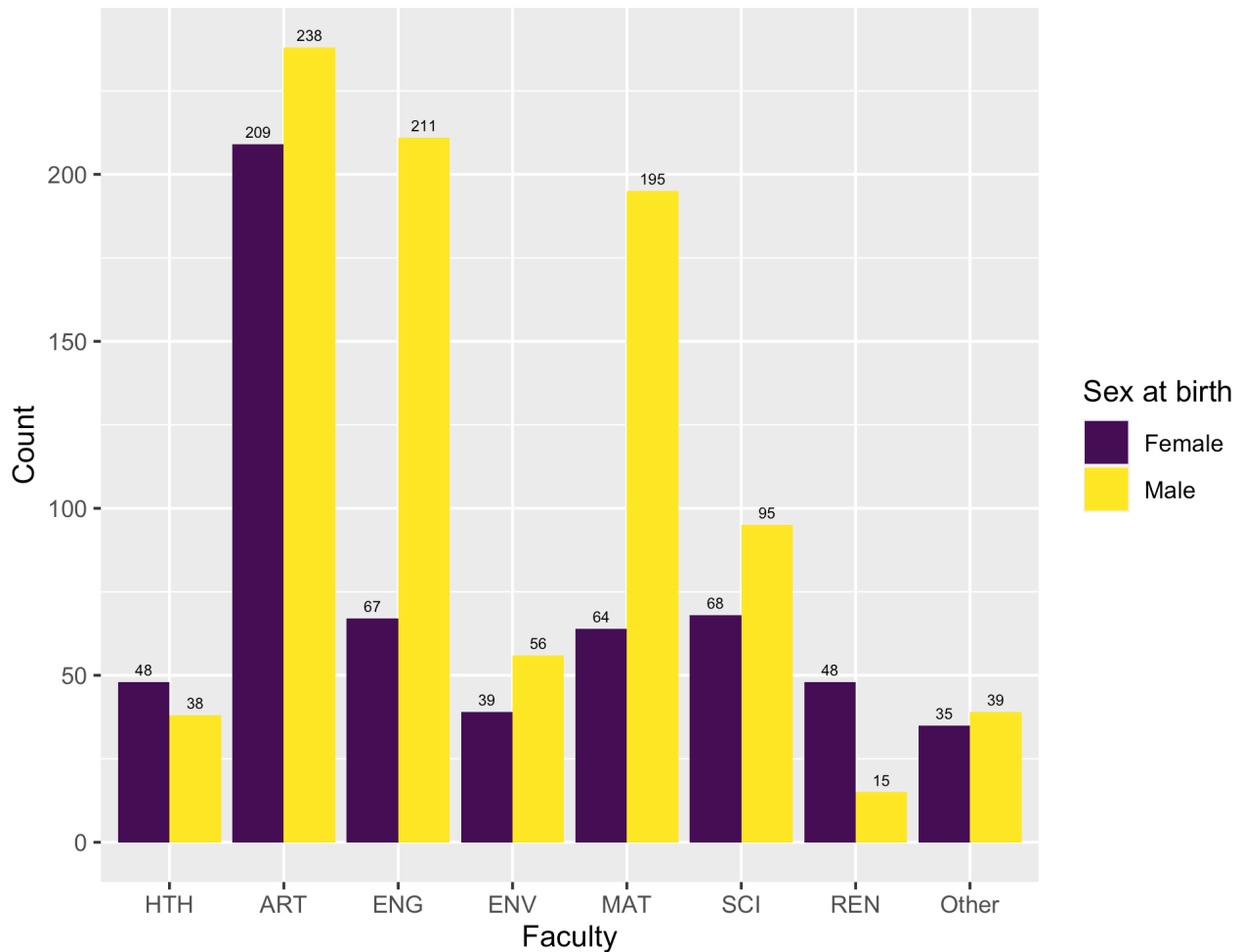


Figure 7: Number of Winter 2024 instructors by Faculty and Sex.

### Instructor Sex by Appointment Type

Figure 8 further breaks down our sample characteristics by Winter 2024 instructor appointment type and sex at birth. In general, male instructors outnumber female instructors across all instructor types. Nearly four times more Winter 2024 Professors were male (168 males, 43 females) and two times more Winter 2024 Associate Professors were male (120 males, 58 females). The makeup of Winter 2024 instructors in remaining instructor types was less drastically imbalanced: Assistant Professors (63 males, 48 females), Lecturers (92 males, 58 females), and Sessional (200 males, 153 females).

## An Analysis of Winter 2024 Student Course Perceptions Survey Responses

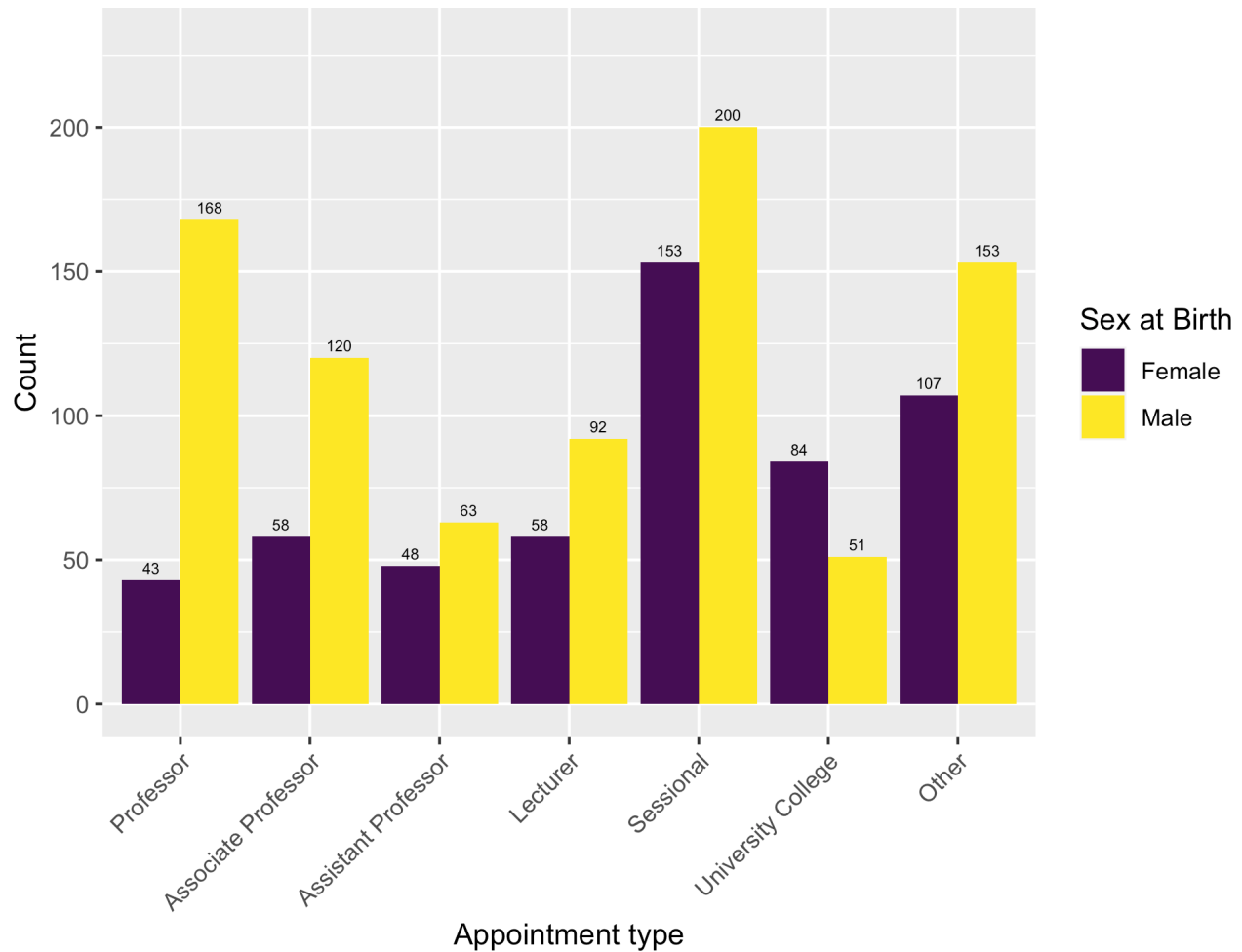


Figure 8: Number of Winter 2024 instructors by appointment type and sex at birth.

### Course-level Variable: Class Size

Figure 9 outlines the number of Winter 2024 courses included in this study by class size. The bulk of courses (1,994) had 100 or fewer students; only a small number (62) had 201+ students enrolled.

## An Analysis of Winter 2024 Student Course Perceptions Survey Responses

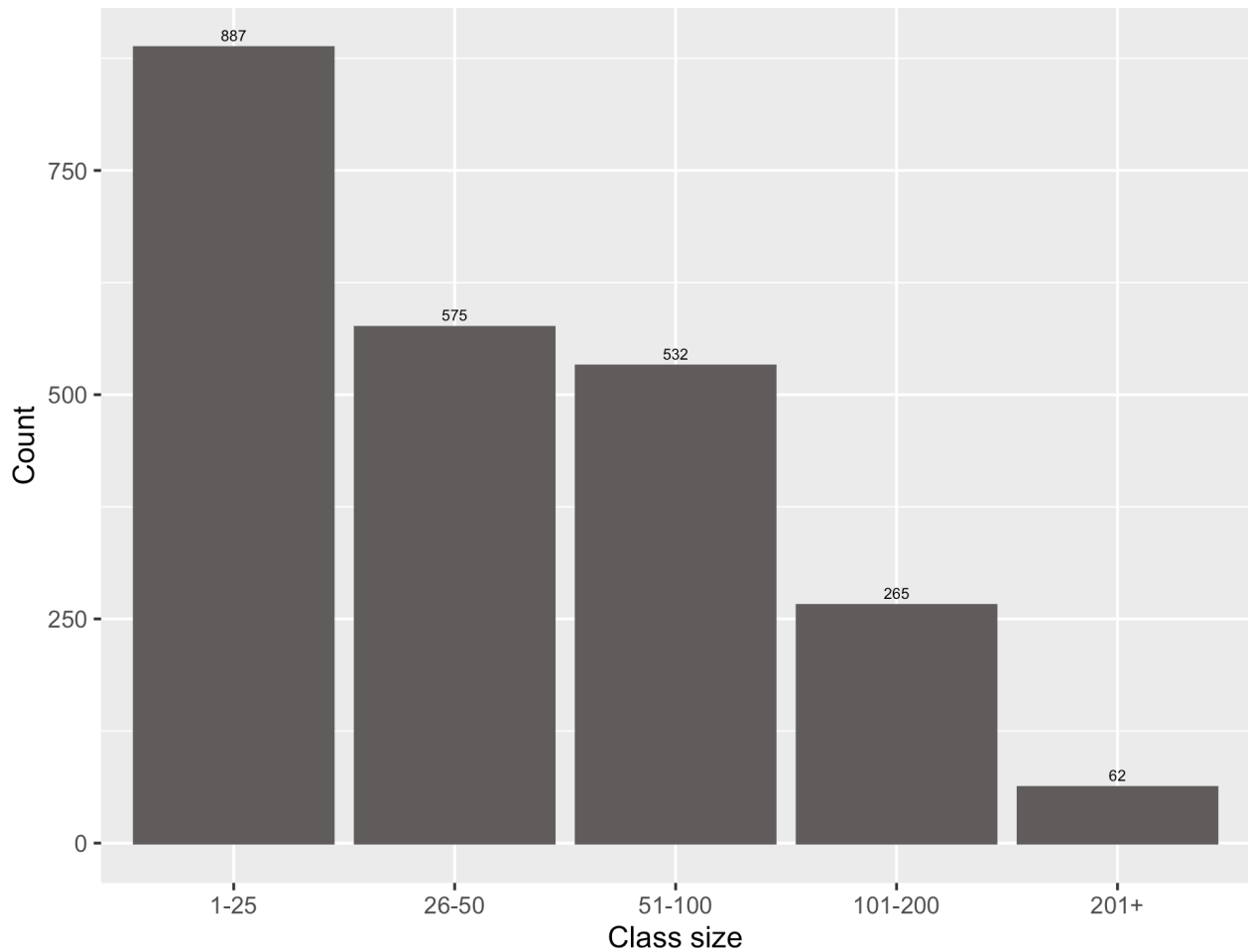


Figure 9: Number of Winter 2024 courses by class size.

### Course-level Variable: Course Type by Faculty

The majority of Winter 2024 courses included in this study were held in person, as reflected in Figure 10. The Faculty of Arts had the highest number of online course offerings (103), but still held just under seven times that number of courses in person (694). Engineering held the second highest number of in-person courses (441), with only 19 courses offered online and Science held the least online courses (8) in contrast to 234 in person.

## An Analysis of Winter 2024 Student Course Perceptions Survey Responses

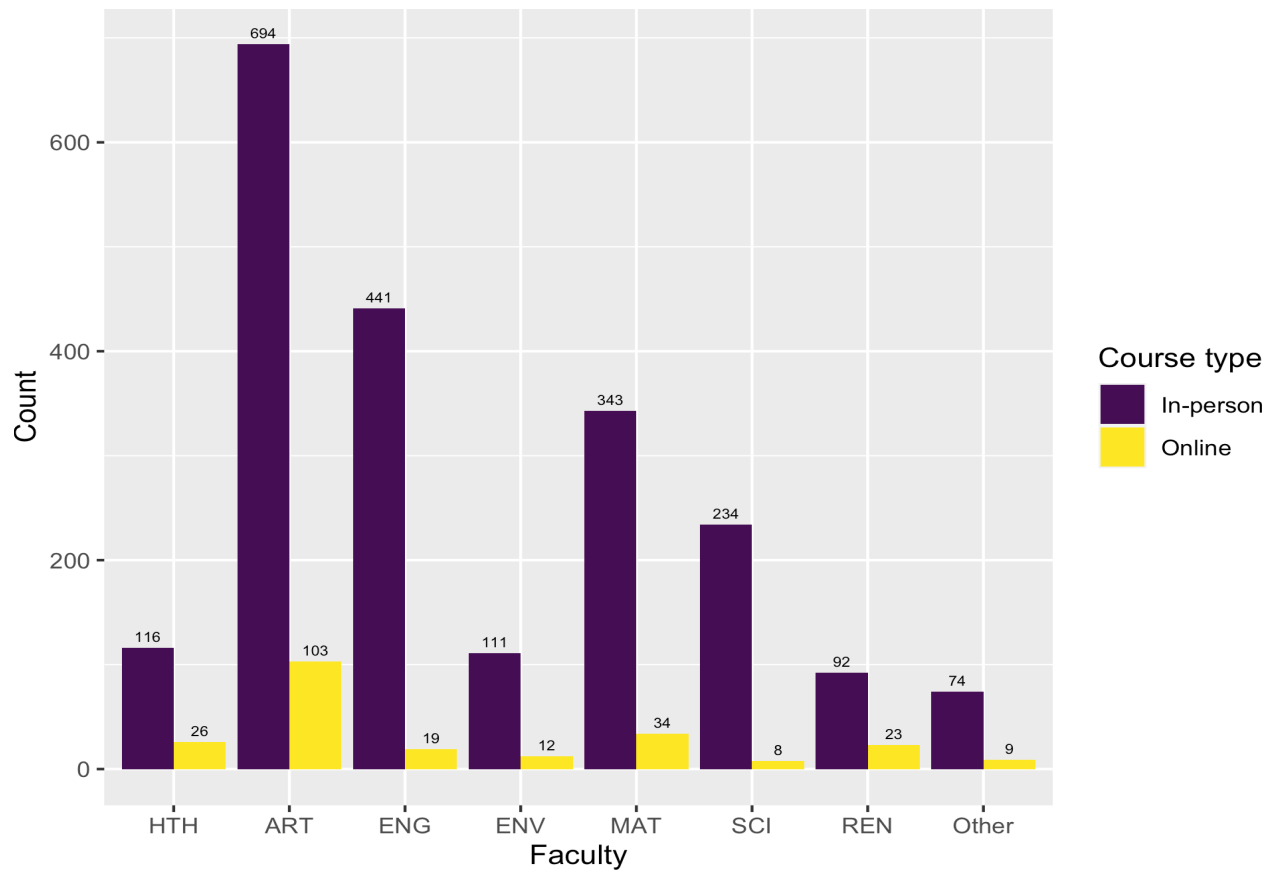


Figure 10: Number of Winter 2024 course sections by faculty and course type.

### SCP Responses by Response Item

Figure 11 displays the frequency with which student respondents assigned each rating for the six core questions on the SCP. The ‘No Basis for Rating’ response was selected only 1-2% of the time. As is often the case with course evaluation surveys, the data is highly skewed, evidenced by the clustering of scores at the high-end of the scale. Specifically, respondents selected ‘Agree’ and ‘Strongly Agree’ between 71% and 82% of the time across all six items. Conversely, respondents selected ‘Disagree’ and ‘Strongly Disagree’ between 6% and 12% of the time across all six items. We can see that student respondents selected the mid-point (‘Neutral’) option on the five-point scale between 10% and 14% of the time across all six items.

## An Analysis of Winter 2024 Student Course Perceptions Survey Responses

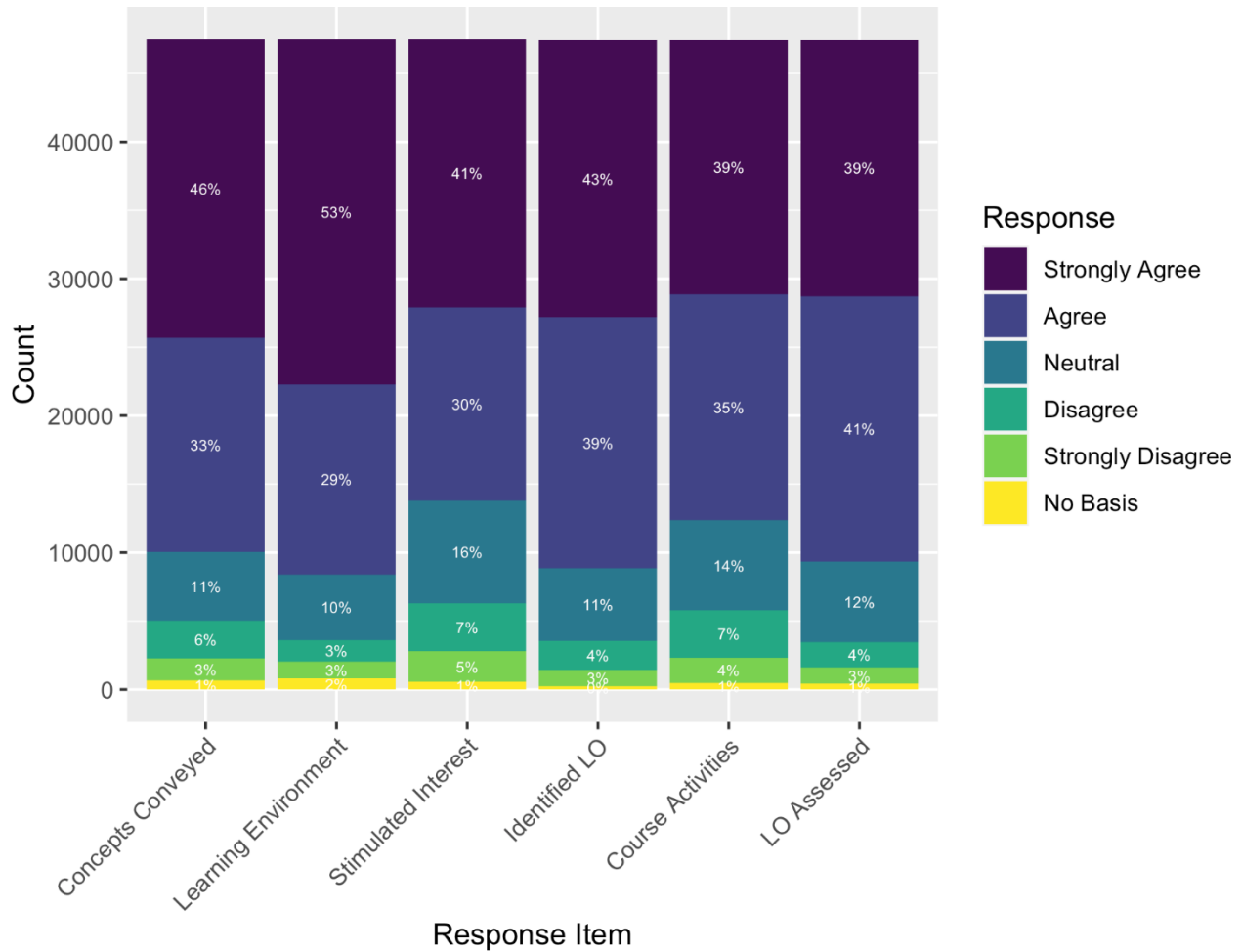


Figure 11: Proportion of Winter 2024 responses per response item.

## SCP Mean Ratings

The following sections contain analyses of mean ratings (MR) of SCP scores for each response item and for various combinations of instructor-level and course-level variables. For a given combination of variables (e.g., assistant professors teaching online courses), the MR in that group is calculated using a mixed-effects model described in detail in Appendix B. Roughly speaking, the calculation is as follows:

- First, we assign a numerical value of 1-5 for each of the response categories from Strongly Disagree (score = 1) to Strongly Agree (score = 5).
- Next, the responses are averaged across students in each instructor/course-section pair. Thus, a multi-section course would have multiple pairs, as would a single section taught by multiple instructors. Students who leave a response item blank or respond with "No Basis for Rating" are not included in the average.
- Using a mixed-effects model approach a course-section-level random effect was identified. Course-section-level effects capture factors specific to the course section offering such as physical comfort of the classroom, student preference for the day or time

of class offering, or other factors impacting student learning experience that are not related to teaching effectiveness. A score for each instructor is then obtained by assuming that the instructor/course-section average is the sum of the instructor score and the course-section-level effect and using the mixed-effects model to extract the former.

- Finally, the instructors' MR in the given group is the average of individual instructor scores across all instructors in the group.

The mixed-effects model also provides standard errors (SE) and 95% confidence intervals (CI) for the MR which account for the number of instructors, the number of course-sections taught by each instructor, and the number of student responses in each instructor/course-section pair.

### Mean Ratings by Response Item and Class Size

Figure 12 displays the mean rating for each of the six response items accounting for class size. Reiterating what we observed in Figure 11, we can see that scores tend to cluster at the higher-end of the five-point scale, with most response items receiving scores between 3.8 and 4.5.

Students tend to rate their learning experiences more positively in smaller classes, as evidenced by higher ratings across all six survey items. This is not unexpected since smaller class sizes have more opportunity for individual student attention, and it aligns with the broader literature as well as past Waterloo SCP analyses. It also is interesting to observe very similar trends in the plotting of the data across all class sizes (the dotted line patterns). We see a peak for Learning\_Environment across all class sizes (ranging from 3.8 in 200+ courses to close to 4.6 in courses with 25 or fewer students). Conversely, Stimulated\_Interest (ranging from 3.6 for 200+ students to 4.3 for 1-25 students) and Course\_Activities (ranging from 3.8 for 200+ students to 4.25 for 1-25 students) receive the lowest mean rating across all class sizes.

## An Analysis of Winter 2024 Student Course Perceptions Survey Responses

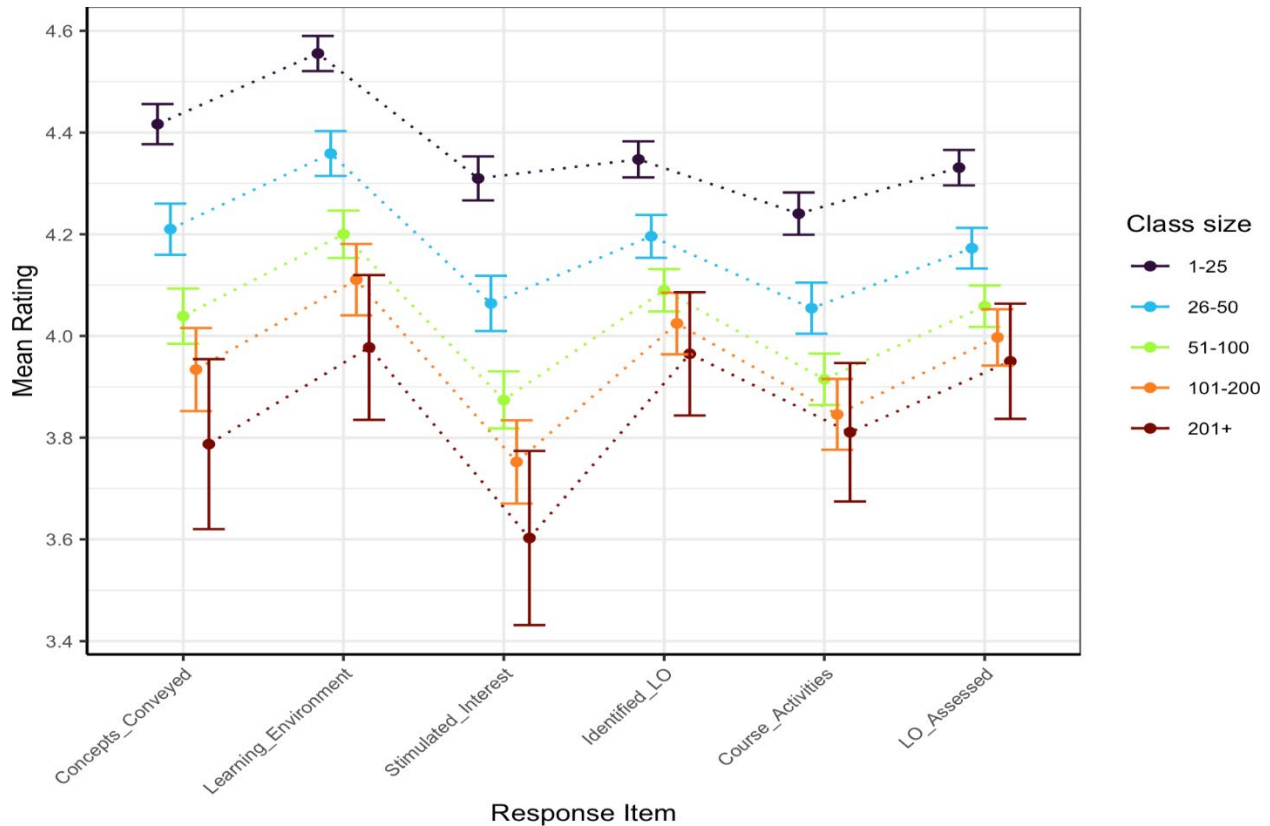


Figure 12: Mean Winter 2024 rating by response item and class size. Error bars correspond to 95% confidence intervals.

### Mean Ratings by Response Item and Faculty of Course Offering

Figure 13 shows the mean ratings across the six core items by Faculty. Perhaps unsurprisingly, results here look similar to what we have observed throughout this report: in general, scores cluster at the higher end of the five-point scale accounting for Faculty (from 3.9 at the lowest end to about 4.5 at the higher end). We can see that across most Faculties Concepts\_Conveyed, Learning\_Environment, Identified\_LOs and LOs\_Assessed received higher scores while Stimulated\_Interest and Course\_Activities received slightly lower scores accounting for Faculty. It is important to emphasize that the differences in practical terms, on the 5-point Likert scale, are very small. For example, the mean rating for Course\_Activities in Science was 3.9 and in Arts was 4.1, which translates to a negligible difference – just 0.2 on the 5-point scale.

## An Analysis of Winter 2024 Student Course Perceptions Survey Responses

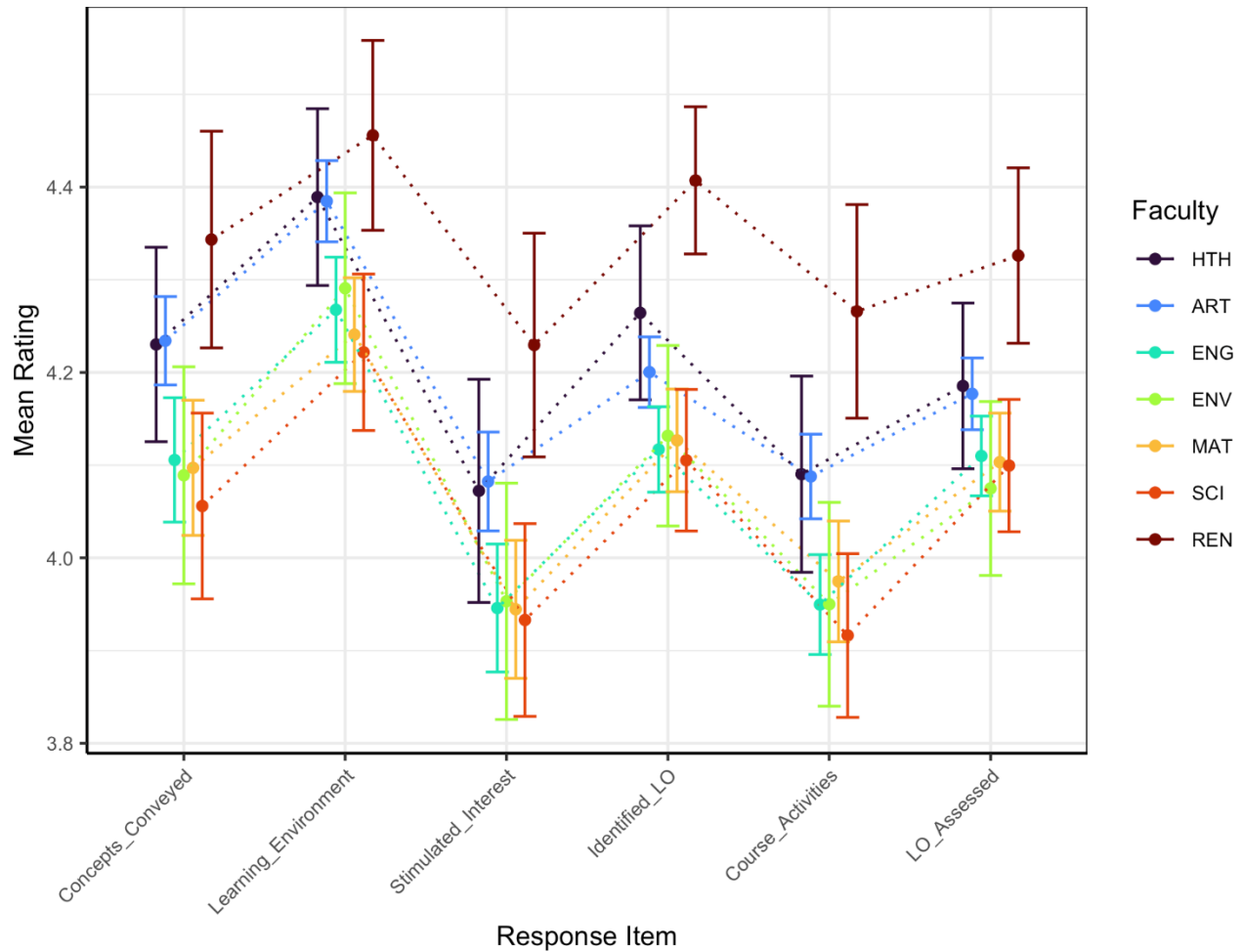


Figure 13: Mean Winter 2024 rating by response item and Faculty of course offering. Error bars correspond to 95% confidence intervals.

## Differences in Mean Ratings

This section examines differences in mean ratings (MR) assigned by students using different instructor- and course-level variables across the six core items on the SCP survey. This includes differences in ratings assigned by students for male instructors and female instructors, different instructor types, classes of different sizes, and different course types (e.g., online or in-person). It is important to note the following analyses are exploratory and do not support causal claims about the effect of sex on SCP ratings. However, they offer an important first step in identifying potential systemic patterns and biases in course evaluations at the University of Waterloo.

Interpreting the Plots:

- Compare plotted points to the zero line
  - The closer to the zero line, the smaller the difference in MR (a difference of 0 suggests no association).
- Compare error bars (which represent 95% confidence intervals) to the zero line:

## An Analysis of Winter 2024 Student Course Perceptions Survey Responses

- When the confidence interval doesn't cross the zero line, it indicates a statistically significant difference in MR using a 95% confidence interval ( $p \leq 0.05$ )
  - When the confidence interval crosses the zero line, it indicates a difference in MR that is not statistically significant using a 95% confidence interval ( $p > 0.05$ )
- The confidence intervals show the difference in scores assigned by students to course surveys taught by those who experience systemic privilege (male) and those who experience systemic oppression (female). The position of the plot point relative to the zero line shows who experiences a benefit from the difference in scores:
  - When the plot point is above the zero line, the difference benefits those who experience systemic privilege.
  - When the plot point is below the zero line, the difference benefits those who experience systemic oppression.
- In cases where a certain combination of covariates does not exist (e.g., assistant professor + class size of 201 + Environment), no data will be displayed in the accompanying figure.
- To protect instructor anonymity, in cases where a certain combination of covariates exists but results in five or fewer instructor-section pairs (e.g., associate professor + class size of 1-25 + Environment), the box is displayed but will be blank.

### Analysis of Differences: Instructor Sex

The sex-based difference in ratings is calculated as follows:

$$MR \text{ assigned to male instructors} - MR \text{ assigned to female instructors}$$

### Difference in MR for Male and Female Instructors

Figure 15 shows the difference in mean ratings assigned by student respondents to male and female instructors across all response items. In all instances, differences observed are not statistically significant, with the exception of one survey item 'stimulated interest' (as evidenced by the error bar crossing or touching the zero line). Across nearly all cases, the non-significant difference is very small (as evidenced by the data points falling close to the zero line).

Importantly the difference in scores for the single statistically significant survey question 'stimulated\_interest' is less than 0.25 points different on the 5 point Likert scale.

## An Analysis of Winter 2024 Student Course Perceptions Survey Responses

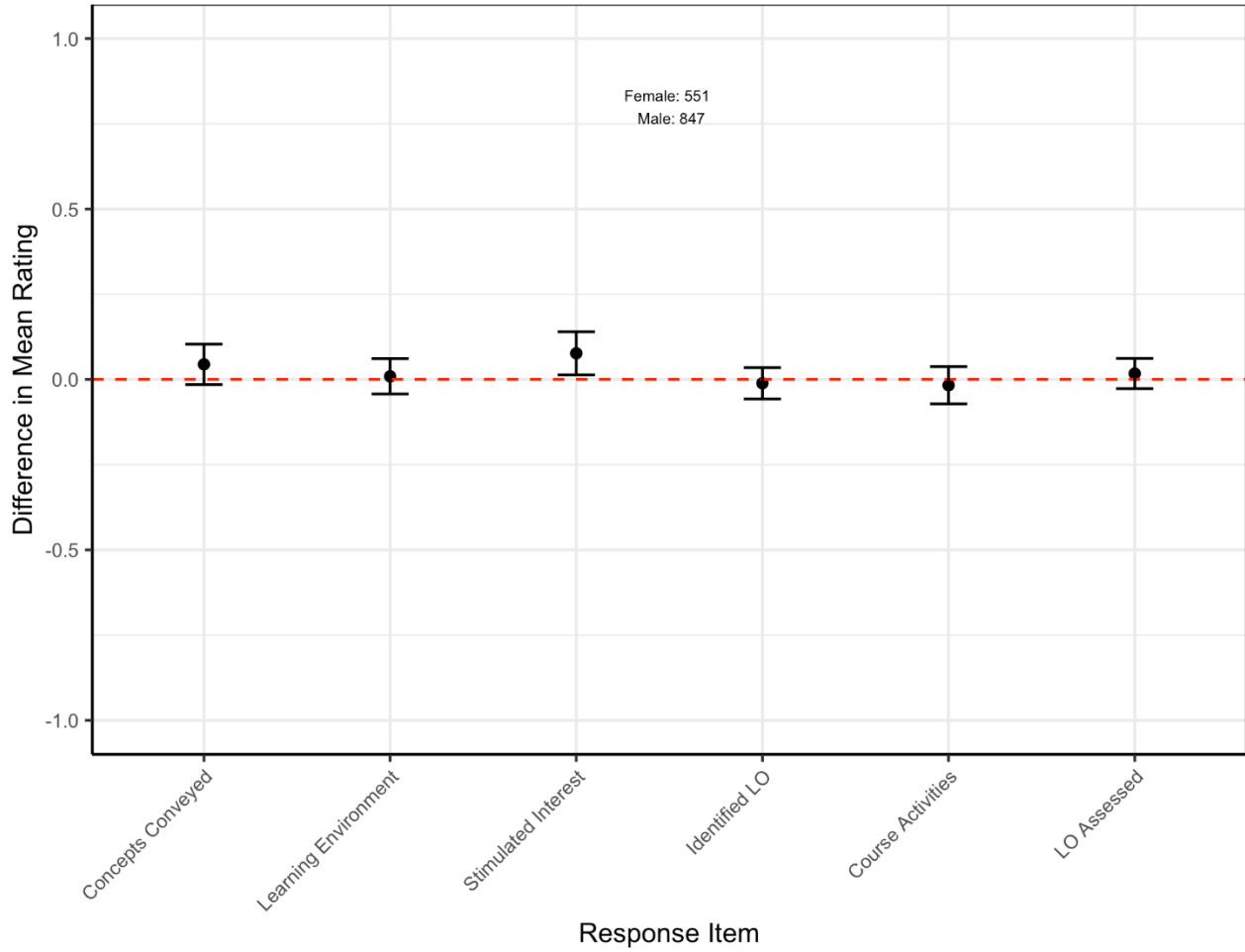


Figure 14: Difference in mean Winter 2024 ratings for male and female instructors per response item. Error bars correspond to 95% confidence intervals.

A clearer picture of the non-significant differences in mean ratings for male and female instructors is evident in Table 3, which also displays the confidence intervals for each response item. Again, across the six core items we can see how the difference in mean ratings on the five-point scale is very small (ranging from -0.02 to 0.08).

Table 1: Difference in mean Winter 2024 ratings for male and female instructors per response item, with 95% confidence interval (C.I.).

|                      | Male Avg. | Female Avg. | Difference | Male C.I.    | Female C.I.  | Difference C.I. |
|----------------------|-----------|-------------|------------|--------------|--------------|-----------------|
| Concepts Conveyed    | 4.18      | 4.14        | 0.04       | (4.15, 4.22) | (4.09, 4.19) | (-0.02, 0.1)    |
| Learning Environment | 4.32      | 4.31        | 0.01       | (4.29, 4.36) | (4.27, 4.36) | (-0.04, 0.06)   |
| Stimulated Interest  | 4.05      | 3.97        | 0.08       | (4.01, 4.09) | (3.92, 4.02) | (0.01, 0.14)    |
| Identified LO        | 4.17      | 4.18        | -0.01      | (4.14, 4.2)  | (4.14, 4.22) | (-0.06, 0.03)   |
| Course Activities    | 4.02      | 4.03        | -0.02      | (3.99, 4.05) | (3.99, 4.08) | (-0.07, 0.04)   |
| LO Assessed          | 4.15      | 4.13        | 0.02       | (4.13, 4.18) | (4.1, 4.17)  | (-0.03, 0.06)   |

## Difference in MR Assigned to Male and Female Instructors by Faculty of Course Offering

We also explored mean ratings assigned by students to male and female instructors accounting for the Faculty of course offering (Figure 16). In all instances except two, differences observed are not statistically significant (as evidenced by the error bar crossing or touching the zero line). And in most cases, the non-significant difference is very small (as evidenced by the data points falling close to the zero line). The only two findings that display statistical significance are for *Stimulated\_Interest* and *Concepts\_Conveyed*, both in Math, where the error bars do not cross the zero line. The data points for this item indicate that student ratings for male instructors for these two questions are higher (about 0.25 or less) than for female instructors. Other than these two response items in Math, we observe a close clustering of scores on the 0 line across remaining Faculties and all response items.

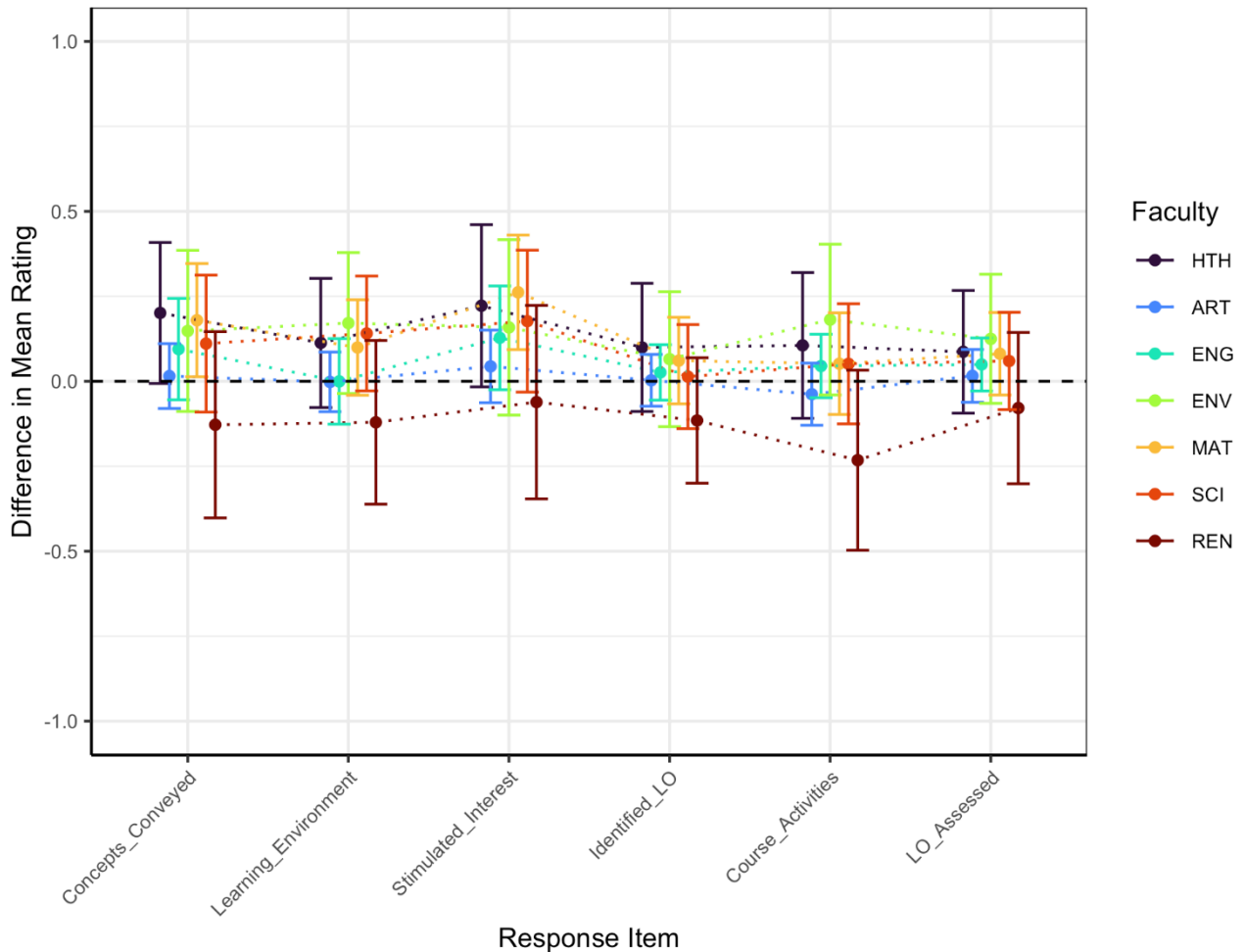


Figure 15: Difference in mean Winter 2024 ratings for male and female instructors by response item and Faculty of course offering. Error bars correspond to 95% confidence intervals.

## Difference in MR for Male and Female Instructors by Faculty of Course Offering and Class Size

We also examined mean ratings at the Faculty level for male and female instructors accounting for class size (Figure 17). For four of the six Faculties (Health, Arts, Environment, and Science) differences observed are not statistically significant (as evidenced by the error bar crossing or touching the zero line). And in most cases, the non-significant difference is very small (as evidenced by the data points falling close to the zero line). However, in Engineering, Math and Science we observe statistically significant differences in scores for males and females for some items. In Engineering in class sizes of 26-50 the following are statistically significant: Concepts\_Conveyed, Stimulated\_Interest, Course\_Activities and LO-Assessed. For each of these items, the error bar is above the zero line indicating males score higher (about 0.5 points or less) than females for these four survey items in classes of this size. In Math we observe statistically significant differences in scores for males and females teaching classes with 101-200 students for only one survey item, LO\_Assessed. The difference in scores for LO\_Assessed is about 0.5 points higher for males teaching in classes with 101-200 students. Finally, in Science we observe a statistically difference in scores for Course\_Activities with female instructors scoring about 0.5 points higher for this item.

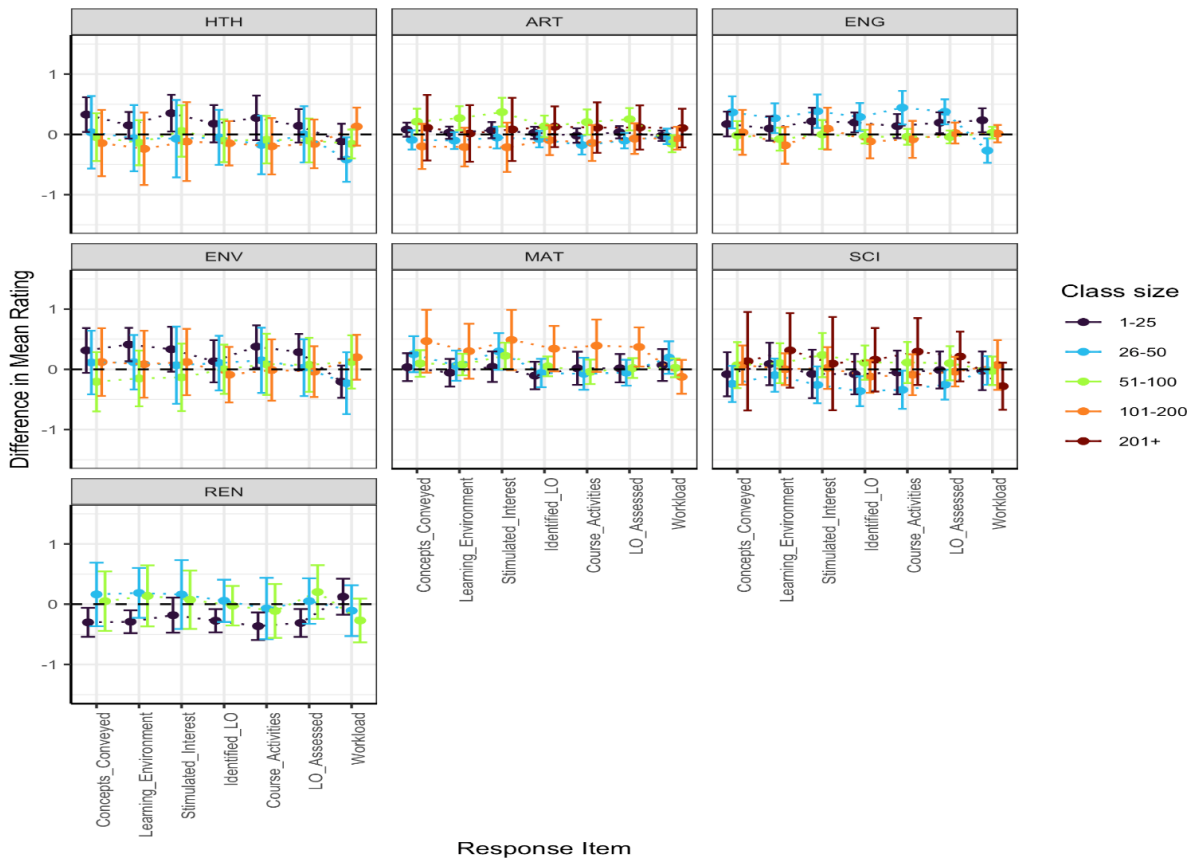


Figure 16: Difference in mean Winter 2024 ratings for male and female instructors by response item, class size, and Faculty of course offering. Error bars correspond to 95% confidence intervals.

## Difference in MR for Male and Female Instructors by Faculty of Course Offering and Course Type

Figure 18 depicts the difference in mean ratings for male and female instructors at the Faculty level accounting for course type (Online or In-person). These results should be considered within the context of the sample sizes. If we refer back to Figure 10, we can see that the majority of courses offered in Winter 2024 were in-person, across all Faculties. In all instances, differences observed are not statistically significant (as evidenced by the error bar crossing or touching the zero line). And in most cases, the non-significant difference is very small (as evidenced by the data points falling close to the zero line).

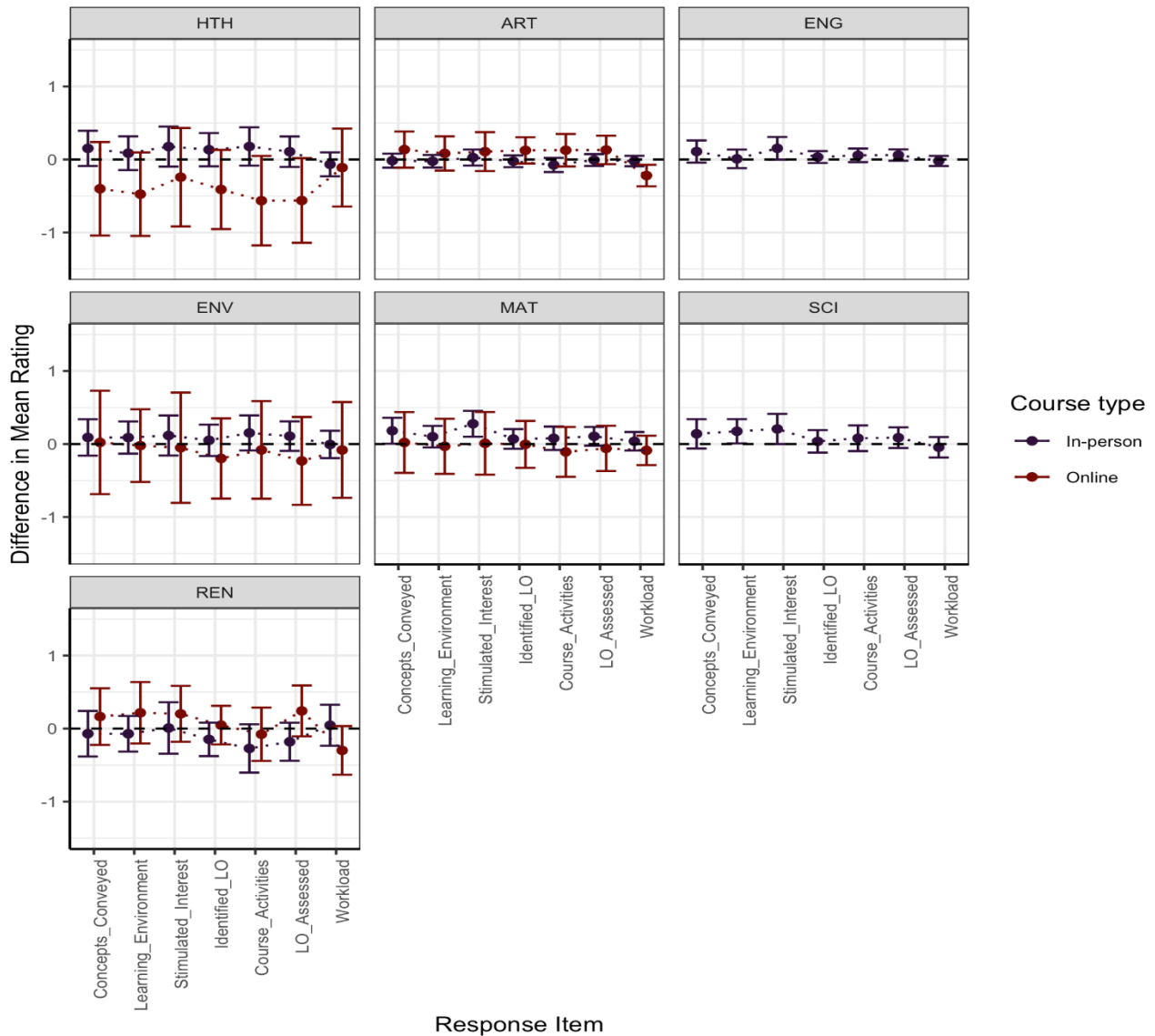


Figure 17: Difference in mean Winter 2024 ratings between male and female instructors per response item by course type and Faculty. Error bars correspond to 95% confidence intervals.

## Difference in MR for Male and Female Instructors by Class Size and Instructor Appointment Type

Figure 17 displays the difference in mean ratings assigned by student respondents to female and male instructors accounting for instructor appointment type and class size. In general, these differences are not statistically significant and the differences in scores is 0 or incredibly close to 0 across all items regardless of course size or professor appointment type.

## An Analysis of Winter 2024 Student Course Perceptions Survey Responses

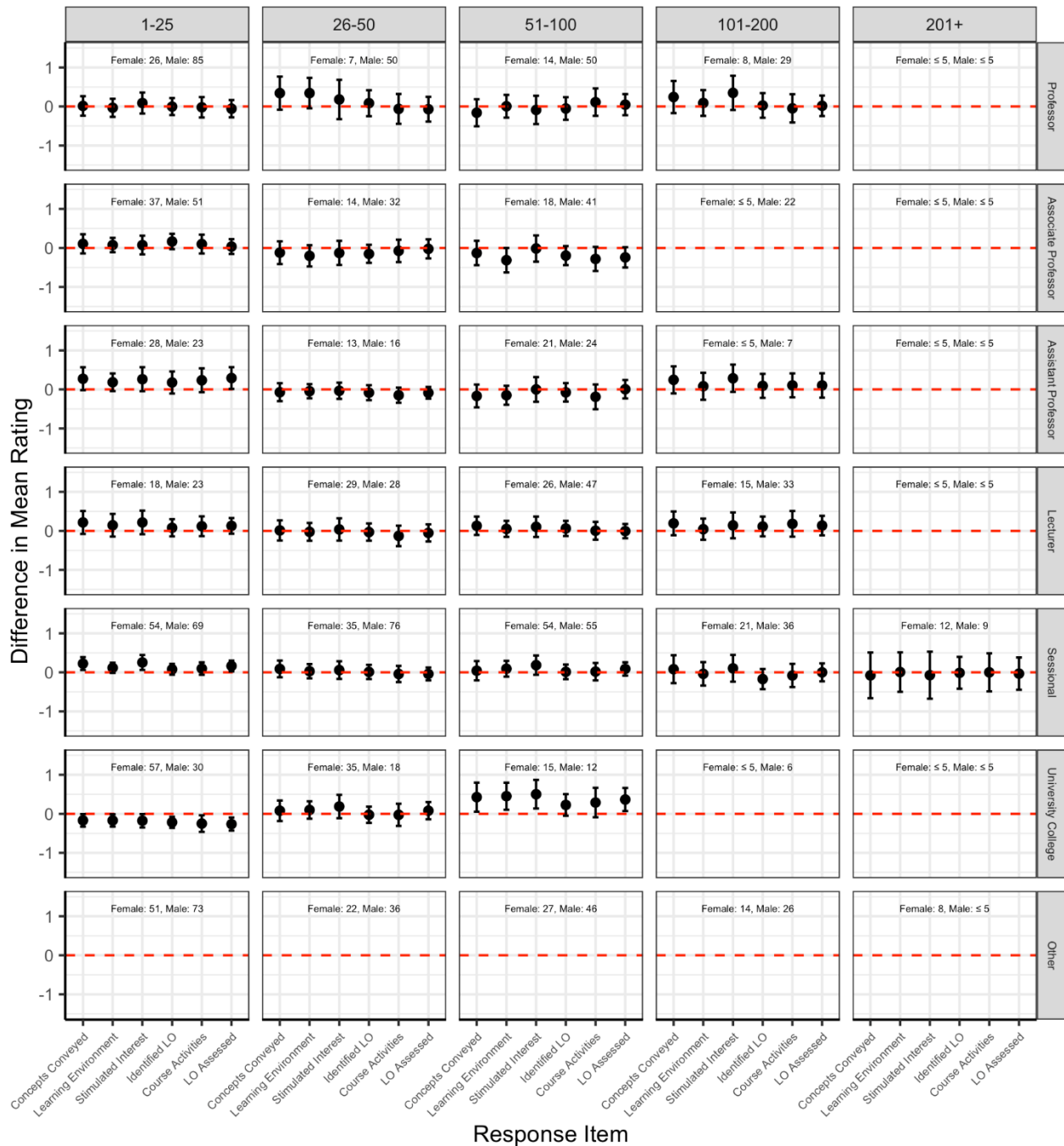


Figure 18: Difference in Winter 2024 mean rating for Male and Female Instructors by Class Size and Instructor Appointment Type. Error bars correspond to 95% confidence intervals.

### Difference in MR for Male and Female Instructors by Class Size

In Figure 17 the difference in mean ratings for male and female instructors across all six survey items is plotted, accounting for class size. It's noteworthy that for larger classes (101-200 and 200+), the sample size is small, which is reflected by the larger confidence intervals; these should be interpreted with extra caution.

## An Analysis of Winter 2024 Student Course Perceptions Survey Responses

In all instances except two, differences observed are not statistically significant (as evidenced by the error bar crossing or touching the zero line). And in most cases, the non-significant difference is very small (as evidenced by the data points falling close to the zero line). The first finding that displays statistical significance is for Concepts\_Conveyed in classes of 1-25 students, where the error bar does not cross the zero line. The data point for this item indicates that student ratings for this question are slightly higher (under 0.1 points) for a class of 1-25 students when they have a male instructor as opposed to a female instructor. The second finding with statistical significance is for Stimulated\_Interest in courses with 1-25 students. For this item, we observe virtually the same difference in scores we see for Concepts\_Conveyed, with males scoring higher (under 0.1 points) than females.

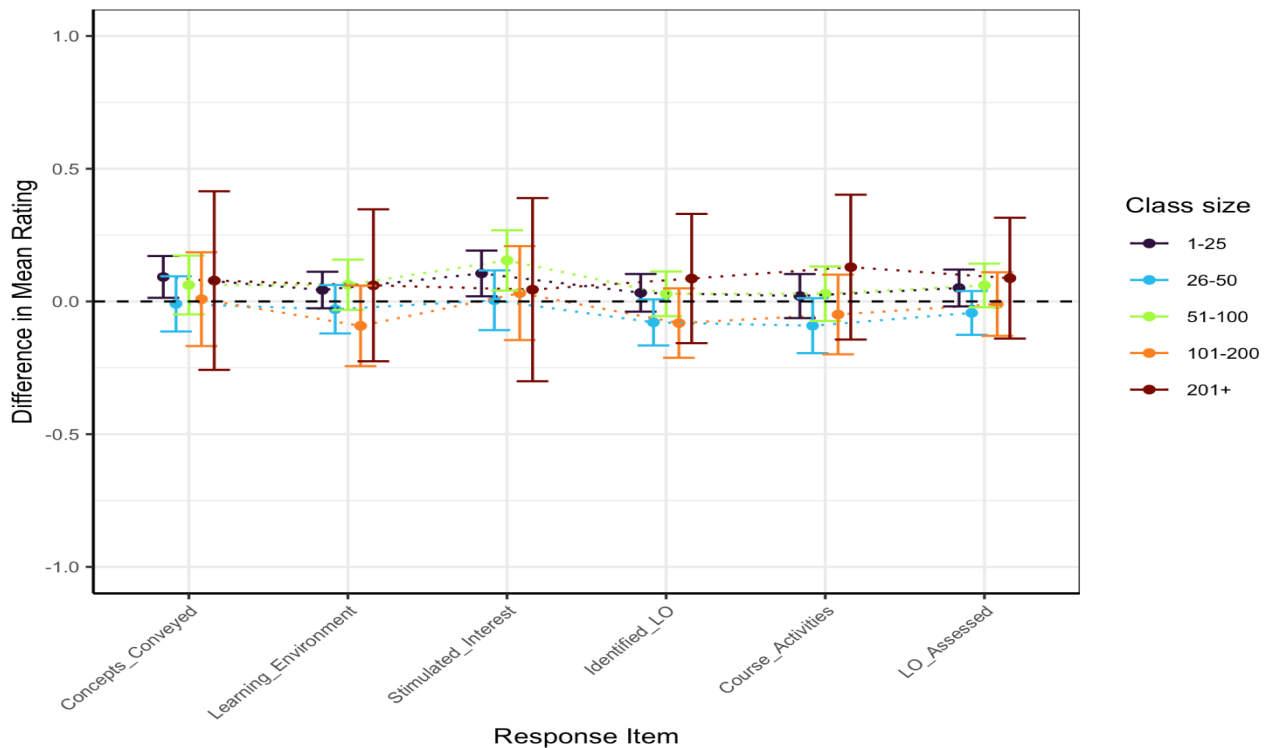


Figure 19: Difference in mean rating for Winter 2024 Male and Female Instructors by class size. Error bars correspond to 95% confidence intervals.

### Difference in MR for Male and Female Instructors by Course Type, and Instructor Appointment Type

Figure 18 shows the difference in mean ratings for female and male instructors, accounting for course type (Online or In-person) and instructor appointment type. In all instances, differences observed are not statistically significant (as evidenced by the error bar crossing or touching the zero line). And in most cases, the non-significant difference is very small (as evidenced by the data points falling close to the zero line).

## An Analysis of Winter 2024 Student Course Perceptions Survey Responses

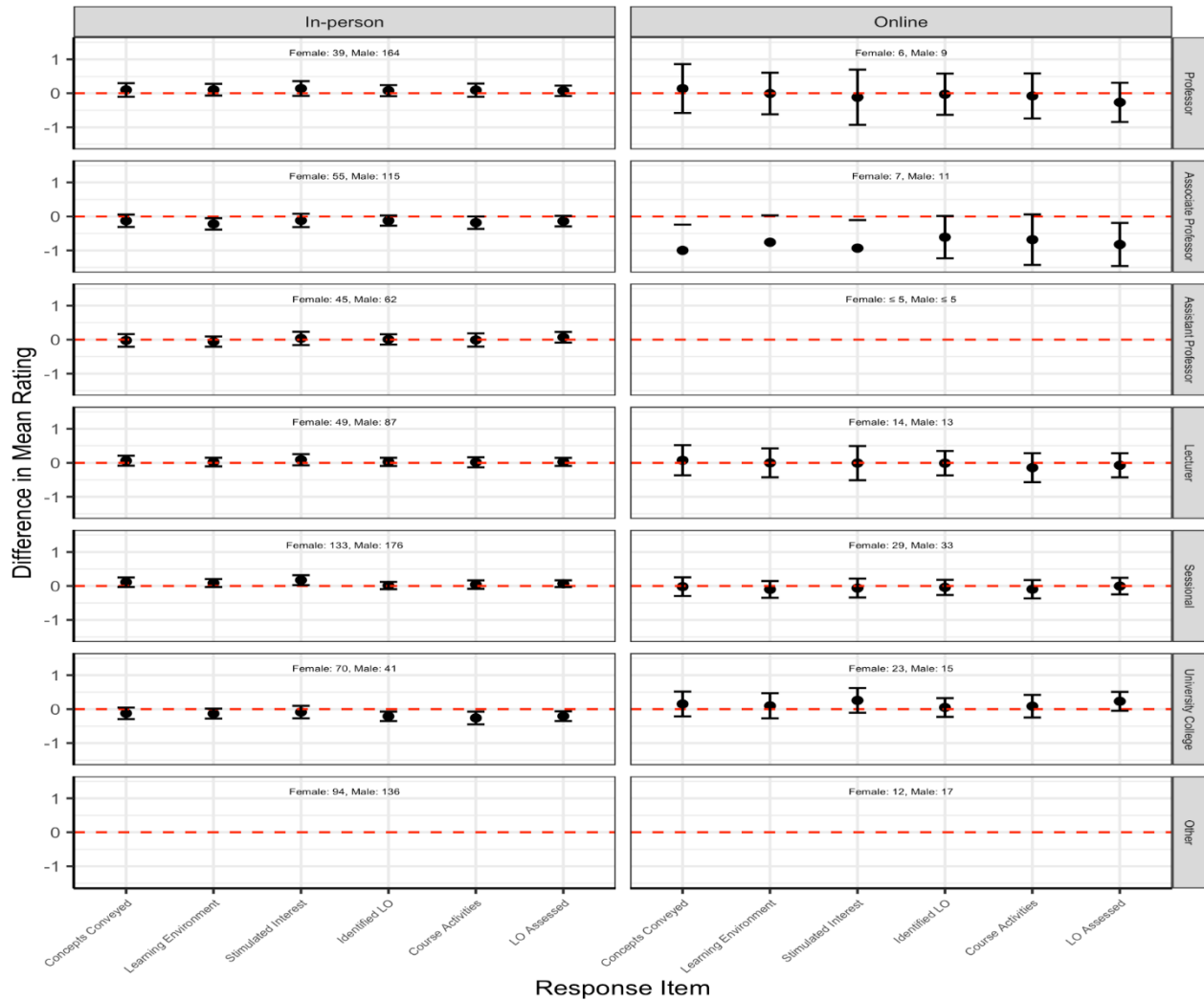


Figure 20: Difference in mean Winter 2024 ratings for male and female instructors by response item, instructor appointment type and course type. Error bars correspond to 95% confidence intervals.

### Difference in MR for Male and Female Instructors by Faculty of Course Offering and Instructor Appointment

At the Faculty level we also examined the difference in mean ratings for male and female instructors accounting for instructor appointment type (Figure 19). In all instances except two, differences observed are not statistically significant (as evidenced by the error bar crossing or touching the zero line). And in most cases, the non-significant difference is very small (as evidenced by the data points falling close to the zero line). We observe significant differences in scores for sessional instructors in Engineering for the following three items: Concepts\_Conveyed, Learning\_Environment and Stimulated\_Interest. The data point for these items indicate that sessional male instructors score slightly higher for these three items (0.5 points or less). In Science, we see statistically significant differences in scores for Professors for Learning\_Environment, Stimulated\_Interest, Course\_Activities, and LO\_Assessed. The data points for these items indicate that male professors score about 0.5 points higher (slightly more

## An Analysis of Winter 2024 Student Course Perceptions Survey Responses

than 0.5 points for Course\_Activities) across these items when compared to female professors in Science.

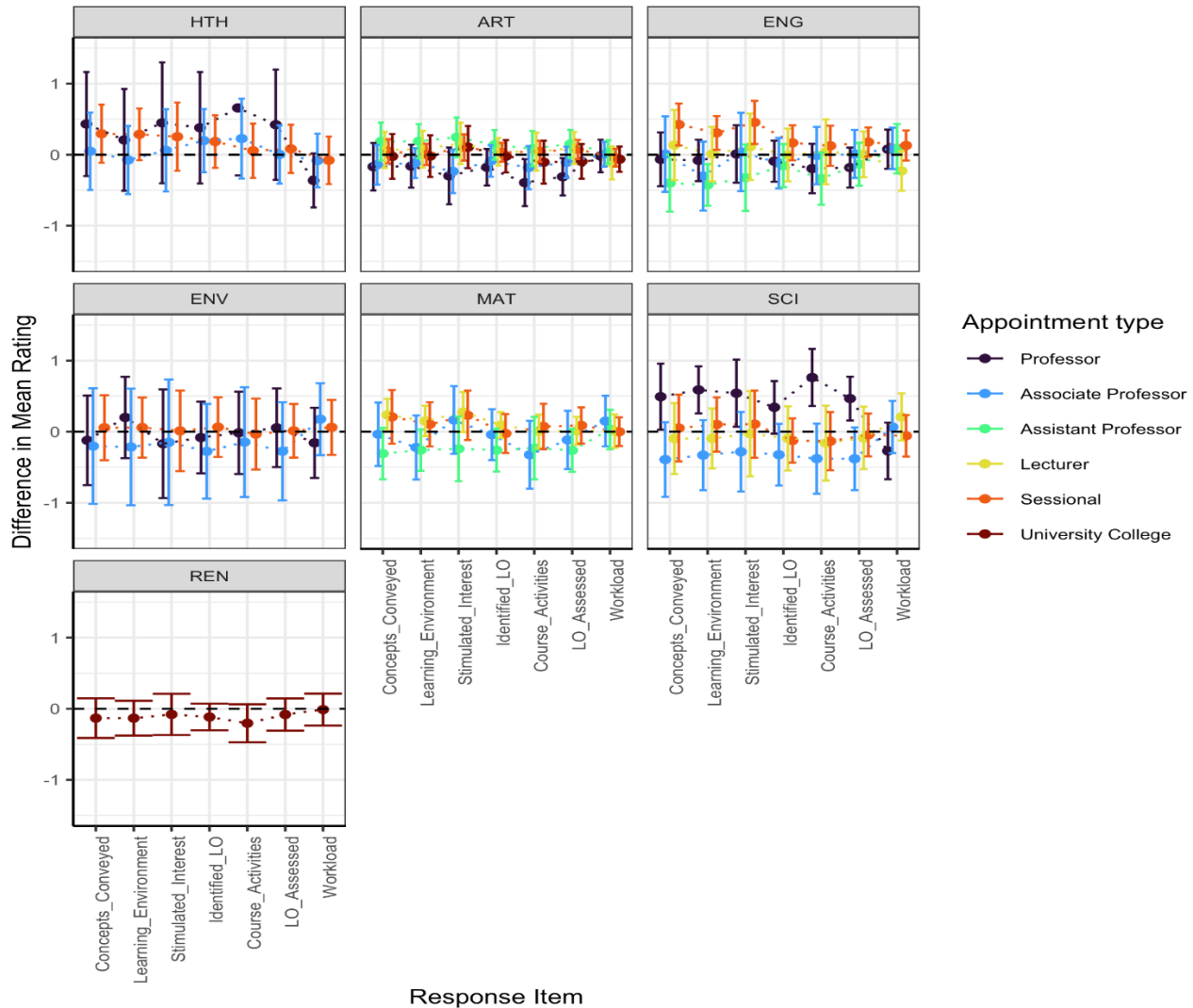


Figure 21: Difference in Winter 2024 mean rating between male and female instructors per response item by instructor appointment and Faculty. Error bars correspond to 95% confidence intervals.

## Difference in MR for Male and Female Instructors by Course Type and Class Size

Figure 20 shows the differences in mean ratings for female and male instructors accounting for course type (Online or In-person) and class size. In all instances, differences observed are not statistically significant (as evidenced by the error bar crossing or touching the zero line). And in most cases, the non-significant difference is very small (as evidenced by the data points falling close to the zero line).

## An Analysis of Winter 2024 Student Course Perceptions Survey Responses

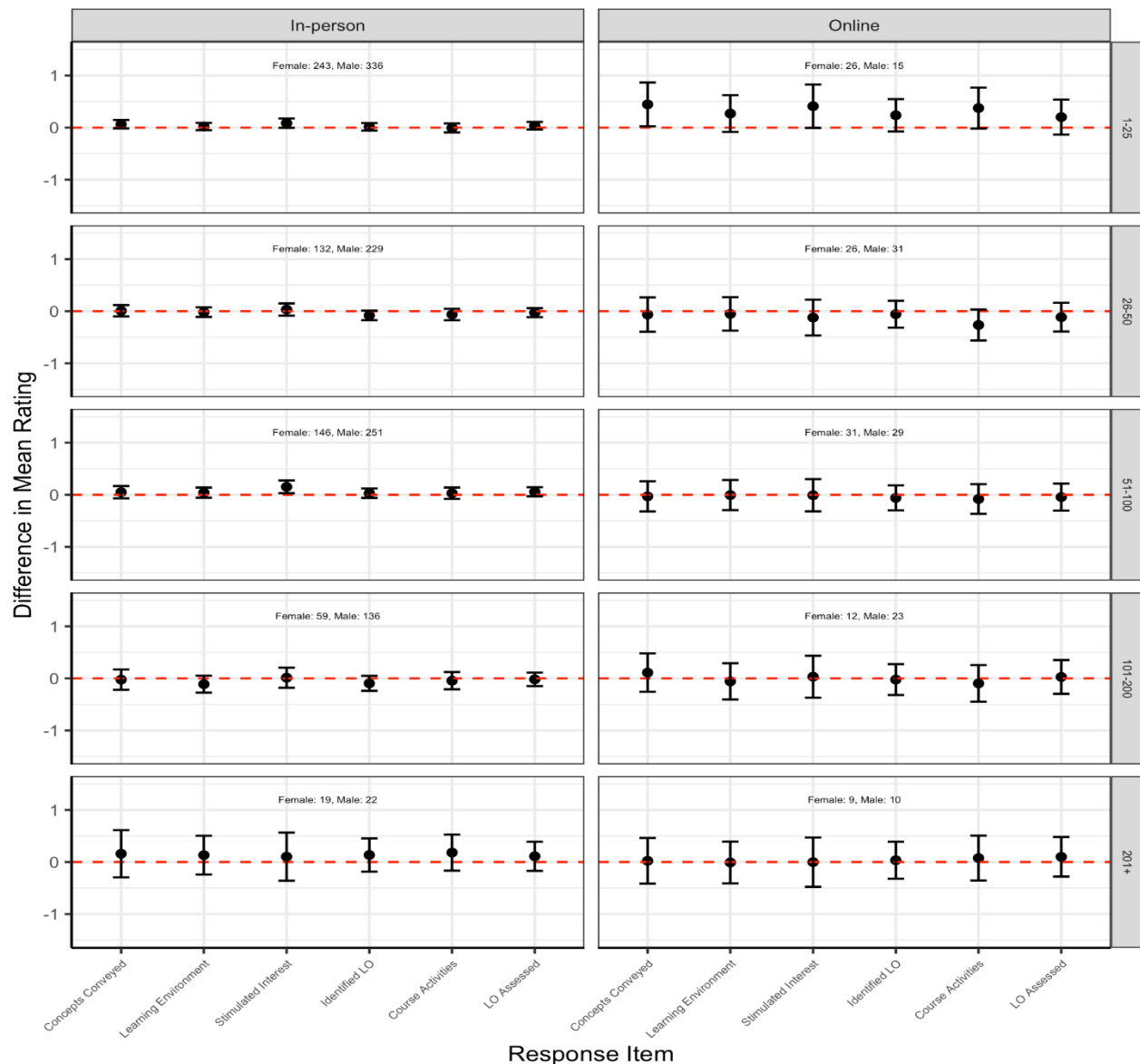


Figure 22: Difference in Winter 2024 mean rating for Male and Female Instructors by Course Type and Class Size. Error bars correspond to 95% confidence intervals.

## Recap of Key Findings

This report presents the third institutional analysis of Student Course Perceptions (SCP) survey data since the introduction of the revised SCP instrument in Winter 2022. Drawing on 47,737 student responses across 2,440 section instructor pairs, the analysis examined the extent to which instructor-level and course-level variables are associated with SCP ratings in Winter 2024.

Overall, the findings closely mirror patterns observed in prior SCP analyses, including Winter 2022 and Winter 2023:

## An Analysis of Winter 2024 Student Course Perceptions Survey Responses

- Across all six core SCP items, student ratings remain strongly skewed toward the positive end of the five-point scale, with the majority of responses falling within the “Agree” and “Strongly Agree” categories. The mid-point (“Neutral”) option was selected infrequently, and the “No Basis for Rating” option accounted for only a very small proportion of responses. This consistency suggests that the SCP continues to function as a stable and interpretable instrument for capturing student perceptions of their learning experiences.
- At the course level, class size remains one of the most consistent correlates of SCP ratings. Smaller classes are associated with slightly higher mean ratings across all six items, while larger classes tend to receive marginally lower ratings. This pattern aligns with both the broader literature on teaching and learning and previous Waterloo SCP analyses. Importantly, however, differences across class sizes are modest in magnitude, and the overall patterning of items remains consistent regardless of class size. Across all class sizes, items related to the learning environment and clarity of learning outcomes tend to receive the highest ratings, while items related to stimulating interest and course activities tend to receive comparatively lower ratings.
- Differences in mean ratings across Faculties are similarly small. While some variation exists, mean ratings across Faculties cluster tightly within a narrow range on the five-point scale. Across most Faculties, students rate concepts conveyed, learning environment, and learning outcomes slightly more positively than stimulated interest and course activities. As with class size, these differences are minimal in practical terms and should not be interpreted as meaningful distinctions in teaching quality across Faculties.
- Instructor-level analyses were constrained by substantial nonresponse to the Equity Survey, particularly for racial identity and Indigenous identity. As a result, analyses examining SCP ratings by racial identity and Indigenous identity could not be conducted for Winter 2024 without risking compromised anonymity or unreliable inference. Available equity data were therefore used descriptively only. Analyses that incorporated instructor time in Canada suggest very small differences in mean ratings across groups, with overlapping confidence intervals and no practically meaningful differences on the five-point scale.
- Analyses examining differences in mean ratings assigned to male and female instructors indicate that, overall, sex-based differences in SCP ratings are small and overwhelmingly not statistically significant. Where statistically significant differences do emerge, typically within specific combinations of Faculty, class size, or appointment type, the magnitude of these differences remains modest, generally well under half a point on the five-point scale. These findings are consistent with previous SCP analyses at Waterloo and reinforce the importance of interpreting isolated statistically significant results with caution, particularly in the context of multiple comparisons and small subgroup sizes.

Taken together, the Winter 2024 findings suggest a high degree of stability in SCP results over time. Observed differences associated with instructor- and course-level characteristics are generally small in magnitude and consistent in direction with prior analyses, underscoring the value of SCP data as one component within a broader, holistic framework for teaching assessment rather than as a standalone or high-stakes evaluative measure.

### Concluding Remarks

The Winter 2024 SCP analysis provides further evidence that the revised Student Course Perceptions survey is functioning as intended: producing stable, interpretable results that reflect broad patterns in student perceptions of teaching and learning while avoiding large or volatile differences associated with instructor or course characteristics. Across three consecutive institutional analyses, SCP ratings have demonstrated remarkable consistency, both in overall distribution and in the relative patterning of items.

Crucially, the small magnitude of observed differences across instructor- and course-level variables highlights a key interpretive challenge—and opportunity—for the institution. On their own, numerical SCP scores provide limited insight into *why* students responded as they did or *how* instructors might meaningfully act on this feedback. Differences of a few tenths of a point on a five-point scale, even when statistically significant, offer little explanatory power without additional context.

It is within this context that open-ended SCP comments become especially important. Student comments provide qualitative detail that can help explain small variations in scores, illuminate patterns that are obscured in aggregate quantitative results, and surface specific aspects of the learning experience that matter most to students. Prior SCP analyses at Waterloo have shown that students frequently use comments to articulate nuanced feedback related to learning environment, engagement, clarity of expectations, and instructional practices, often in ways that closely align with the institution's Teaching Effectiveness Framework. When interpreted alongside SCP ratings, these comments offer a richer and more actionable understanding of teaching and learning than scores alone can provide.

At the same time, this analysis also underscores ongoing challenges related to equity data availability. The high level of nonresponse to the Equity Survey in Winter 2024 limited the scope of analyses related to racial identity and Indigenous identity. Addressing these data limitations will be essential for future work aimed at understanding whether, and how, systemic inequities may shape student evaluations of teaching. Without reliable and representative equity data, the institution's ability to examine these questions rigorously remains constrained.

Looking ahead, the findings from Winter 2024 further support continued investment in a multi-measure approach to teaching assessment, one that integrates SCP data with peer review of teaching, teaching dossiers, and other sources of evidence. Within such a framework, SCP results can play a valuable role in identifying patterns, prompting reflection, and informing professional development, while avoiding inappropriate uses that exceed what the data can reasonably support.

As the University of Waterloo continues to refine its teaching assessment practices, longitudinal SCP analyses such as this one will remain critical. By focusing on trends over time, transparency in methods, and careful interpretation of results, the institution is well positioned to use SCP data responsibly in support of teaching development, equity, and educational excellence.

### Moving Forward: Implications for Teaching Assessment Practice

The findings from the Winter 2024 SCP analysis reinforce the importance of maintaining SCPs as one component within a broader, integrated system of teaching assessment. In particular, they underscore the opportunity to more intentionally foreground student responses to open-ended SCP questions as a critical source of context for interpreting quantitative ratings. While SCP scores signal broad patterns, student comments help explain those patterns, offering insight into the specific teaching practices, course structures, and learning conditions that shape student perceptions.

Within a holistic assessment framework, SCP comments can support more meaningful formative use of evaluation data by informing reflective practice, peer review conversations, and professional development initiatives. They also provide an important safeguard against overinterpretation of small numerical differences by anchoring discussions of teaching effectiveness in students' own words and experiences. Leveraging SCP comments alongside SCP ratings, peer review of teaching, and teaching dossiers enables a more balanced, fair, and development-oriented approach to teaching assessment, one that aligns with institutional commitments to educational quality, equity, and evidence-informed decision-making.

Continued attention to data quality, transparency, and access, particularly with respect to equity-related information and the responsible use of qualitative feedback, will be essential as the University of Waterloo continues to evolve its teaching assessment practices.