

An Analysis of Winter 2025 Student Course Perceptions Survey Responses



ASSOCIATE VICE-PRESIDENT, ACADEMIC | Teaching Assessment
Processes uwaterloo.ca/teaching-assessment-processes

An Analysis of Winter 2025 Student Course Perceptions Survey Responses

Prepared June 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 2025 Student Course Perceptions Survey Responses

Contents

Executive Summary.....	4
Introduction	6
Sources of Data	6
Research Aims	6
Primary Analyses	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: Gender	12
Instructor Gender	12
Instructor Sex	14
Instructor Sex by Faculty	14
Instructor Sex by Appointment Type	15
Course-level Variable: Class Size	16
Course-level Variable: Course Type by Faculty	17
SCP Responses by Response Item.....	18
SCP Mean Ratings.....	19
Mean Ratings by Response Item and Class Size	20
Mean Ratings by Response Item and Faculty of Course Offering	21
Differences in Mean Ratings	22
Analysis of Differences: Instructor Sex	23
Difference in MR for Male and Female Instructors	23
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	27

An Analysis of Winter 2025 Student Course Perceptions Survey Responses

Difference in MR for Male and Female Instructors by Class Size and Instructor Appointment Type	29
Difference in MR for Male and Female Instructors by Class Size.....	30
Difference in MR for Male and Female Instructors by Course Type, and Instructor Appointment Type	31
Difference in MR for Male and Female Instructors by Faculty of Course Offering and Instructor Appointment	32
Difference in MR for Male and Female Instructors by Course Type and Class Size.....	33
Recap of Key Findings.....	35
Concluding Remarks.....	36
Moving Forward: Implications for Teaching Assessment Practice	37

Executive Summary

This report presents the University of Waterloo's fourth institutional analysis of Student Course Perceptions (SCP) survey data following the launch of the revised SCP instrument in Winter 2022. Drawing on 46,302 student responses across 2,483 section-instructor pairs from Winter 2025, 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 2025 SCP mean ratings demonstrate a high degree of consistency with those observed in Winter 2022, Winter 2023, and Winter 2024. 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.

Unfortunately, just as we encountered in Winter 2024, substantial nonresponse to the institutional Equity survey, continues to constrain Instructor-level analyses for Winter 2025. As a result, analyses examining SCP ratings by Indigenous identity, racial identity, and gender identity could not be conducted for Winter 2025 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 2025 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 2025 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 2025. The present study is the fourth 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 2025 SCP survey data and, for instructors whose Winter 2025 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, 743 of the 1,392 instructors in this study did not provide information for the racial identity question on the Equity Survey. Additionally, 52 instructors indicated they '*prefer not to answer*' this particular question. Combined, this represents 56% of instructors with missing equity data for this question, almost identical to the 57% of instructors missing equity data in the Winter 2025 dataset. We continue to surmise that the poor response rate results from: 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 2025 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:

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 2025.
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, and 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, 46,302 SCP surveys from 2483 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:

- 1392 unique instructors
- 1544 unique courses
- 2312 distinct sections

Due to issues 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.

Instructor-level Variable: Indigenous Identity

Ten Winter 2025 instructors identified as Indigenous in the Equity Survey (see Figure 1).

An Analysis of Winter 2025 Student Course Perceptions Survey Responses

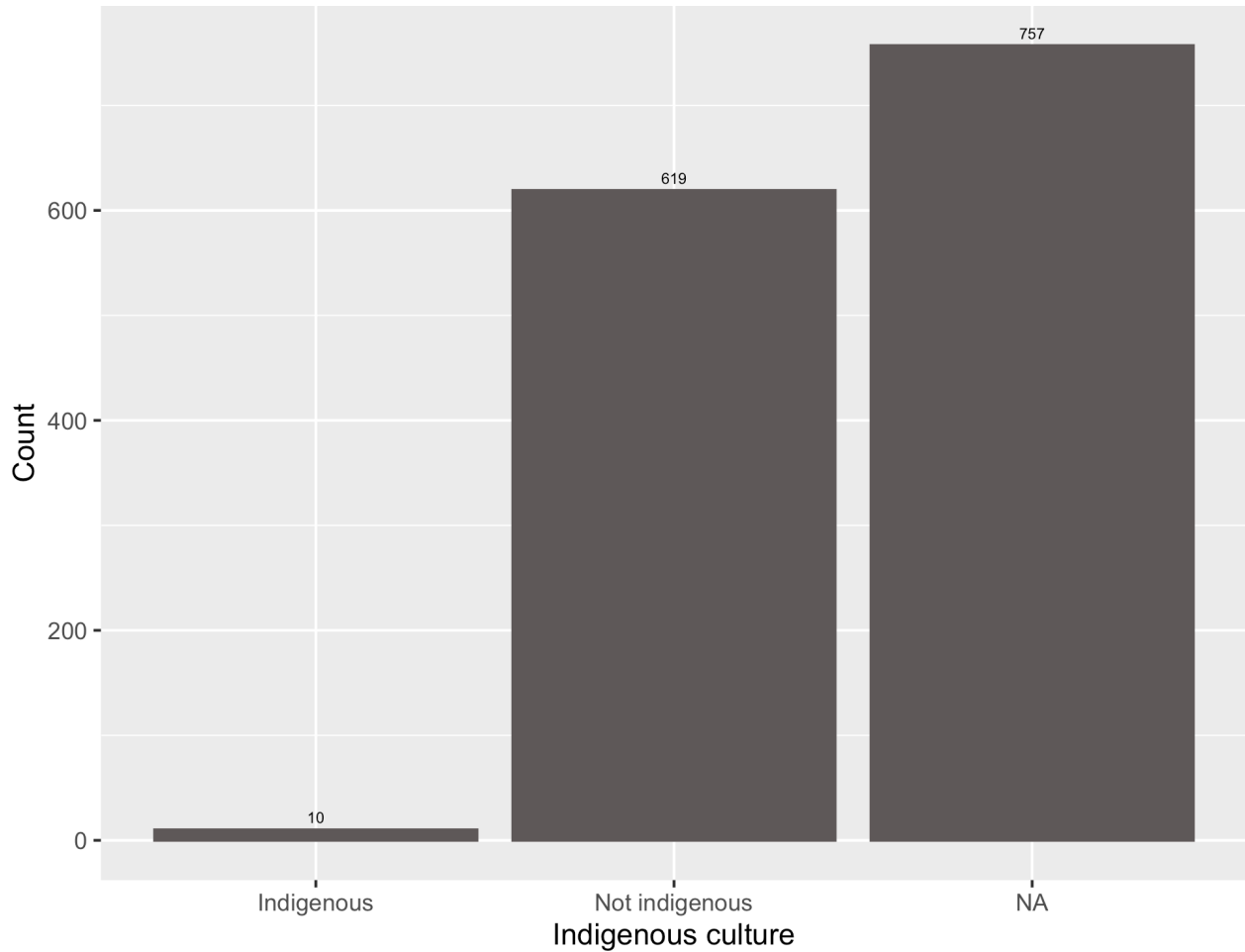


Figure 1: Number of Winter 2025 instructors by self-reported Indigenous identity. Each Equity Survey respondent was asked to select all applicable racial identities. NA represents Winter 2025 instructors who did not complete the Equity Survey.

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 2025 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 744 instructors not providing any racial data and 52 instructors selecting ‘*prefer not to answer*’ for the self-reported racial identity question. Based on the responses we received, the second highest racial category for Winter 2025 instructors are white (402), followed by 57 East Asian, 41 Middle Eastern and 34 instructors indicating they are South Asian. Only 20 instructors who provided racial data indicated they are black.

An Analysis of Winter 2025 Student Course Perceptions Survey Responses

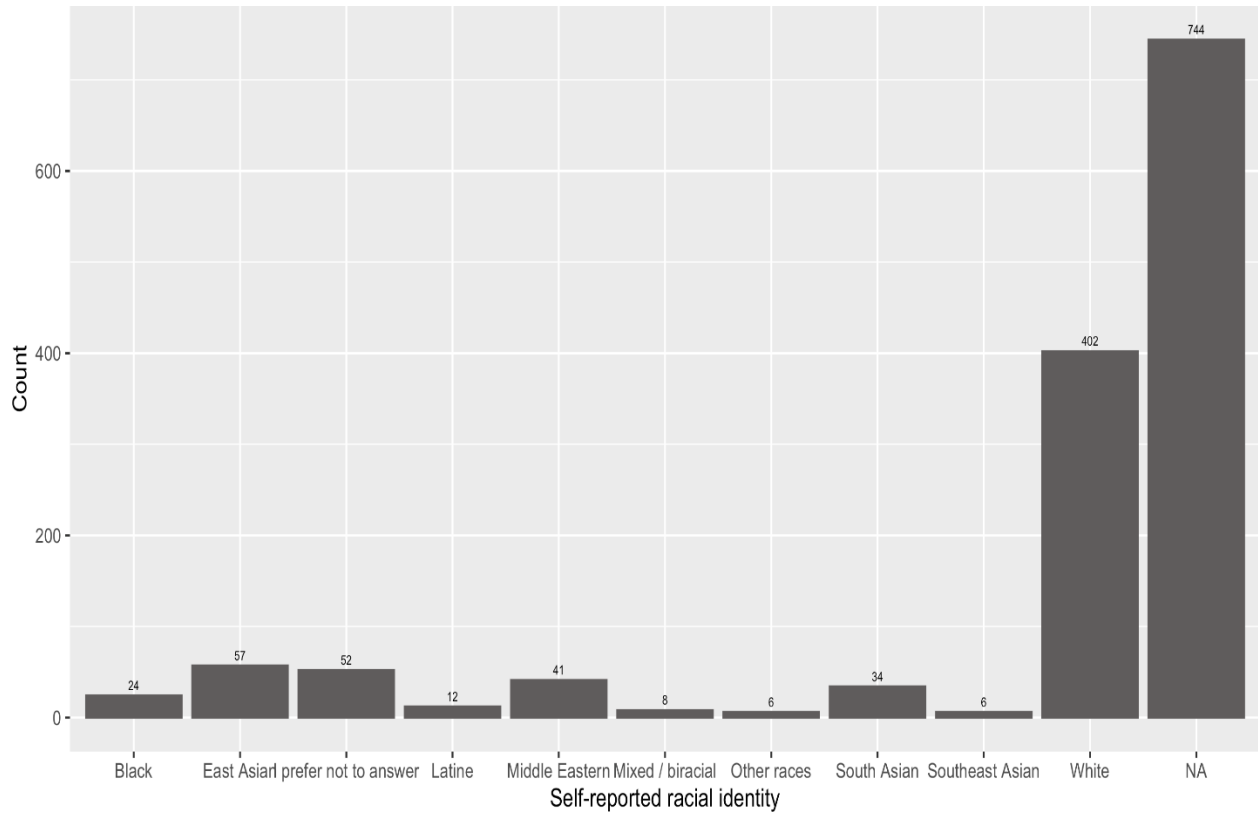


Figure 2: Number of Winter 2025 instructors by self-reported racial identity. Each Equity Survey respondent was asked to select all applicable racial identities. NA represents Winter 2024 instructors who did not provide a response for this item, or did not provide a response.

Racialized Instructors by Faculty

We examined the number of racialized instructors across each Faculty who had SCP surveys in Winter 2025. As shown in Figure 3, there is a high number of instructors who did not report racial data across all six Faculties. Setting the missing cases aside, there were nearly three times more white than racialized instructors teaching in Winter 2025 across three of the six Faculties (Health, Arts, Environment). Just over two times more white than racialized instructors in one Faculty (Science), and almost equal numbers of white and racialized instructors in two Faculties (Engineering, Math).

An Analysis of Winter 2025 Student Course Perceptions Survey Responses

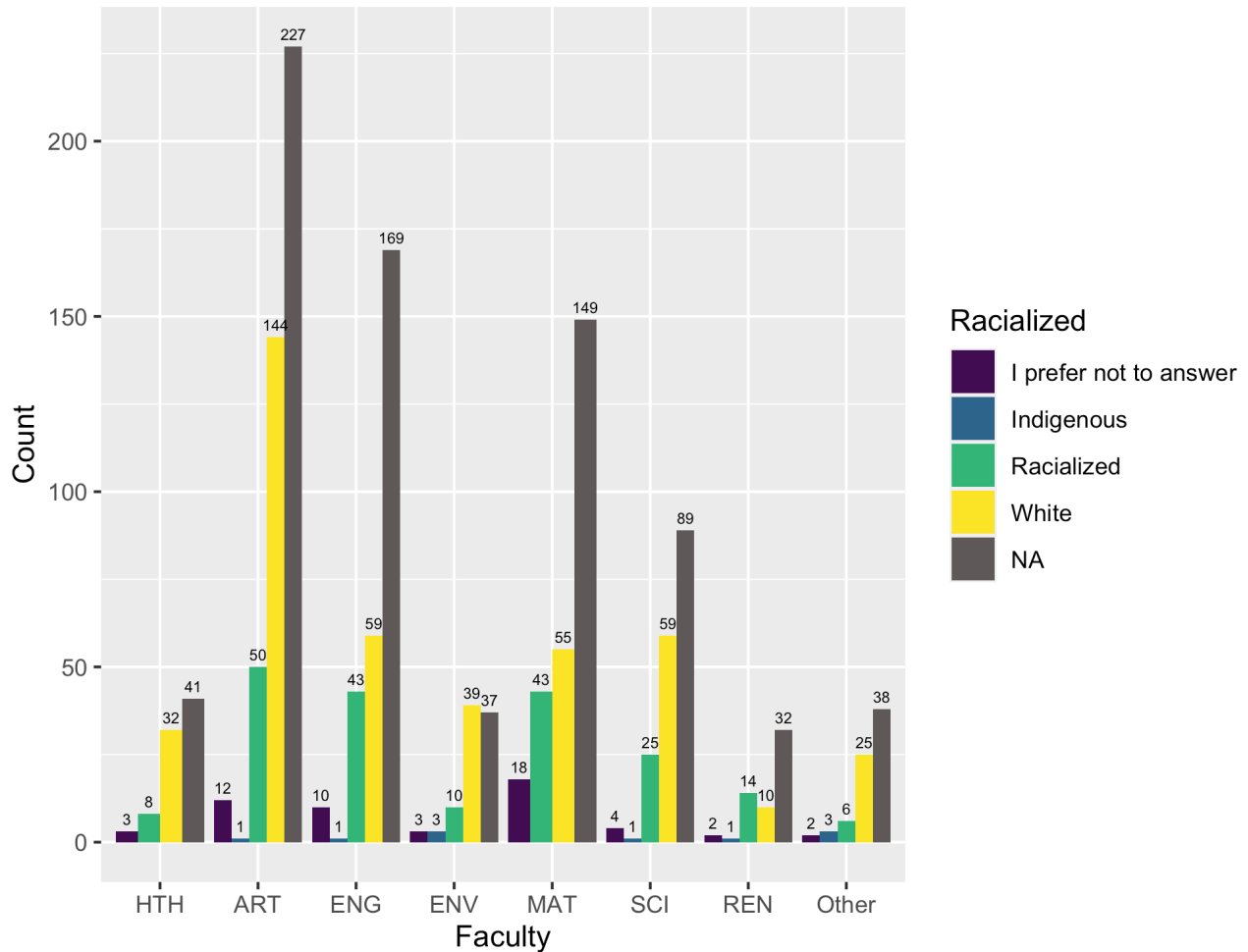


Figure 3: Number of Winter 2025 instructors by faculty and racial identity – Note: NA represents Winter 2025 instructors who did not complete the Equity Survey or preferred not to answer the racial identity question. “Other” includes University Colleges (except for Renison) and those that do not belong to any Faculty (e.g., GRAD).

Instructor Time in Canada

Figure 4 outlines the length of time in Canada for Winter 2025 instructors who responded to the Equity Survey. Overall, 365 instructors in the sample were born in Canada, 236 had been in

An Analysis of Winter 2025 Student Course Perceptions Survey Responses

Canada 5+ years and only 51 instructors reported being in Canada for less than 5 years. Again, we highlight there are 740 instructors who did not provide any data for this question.

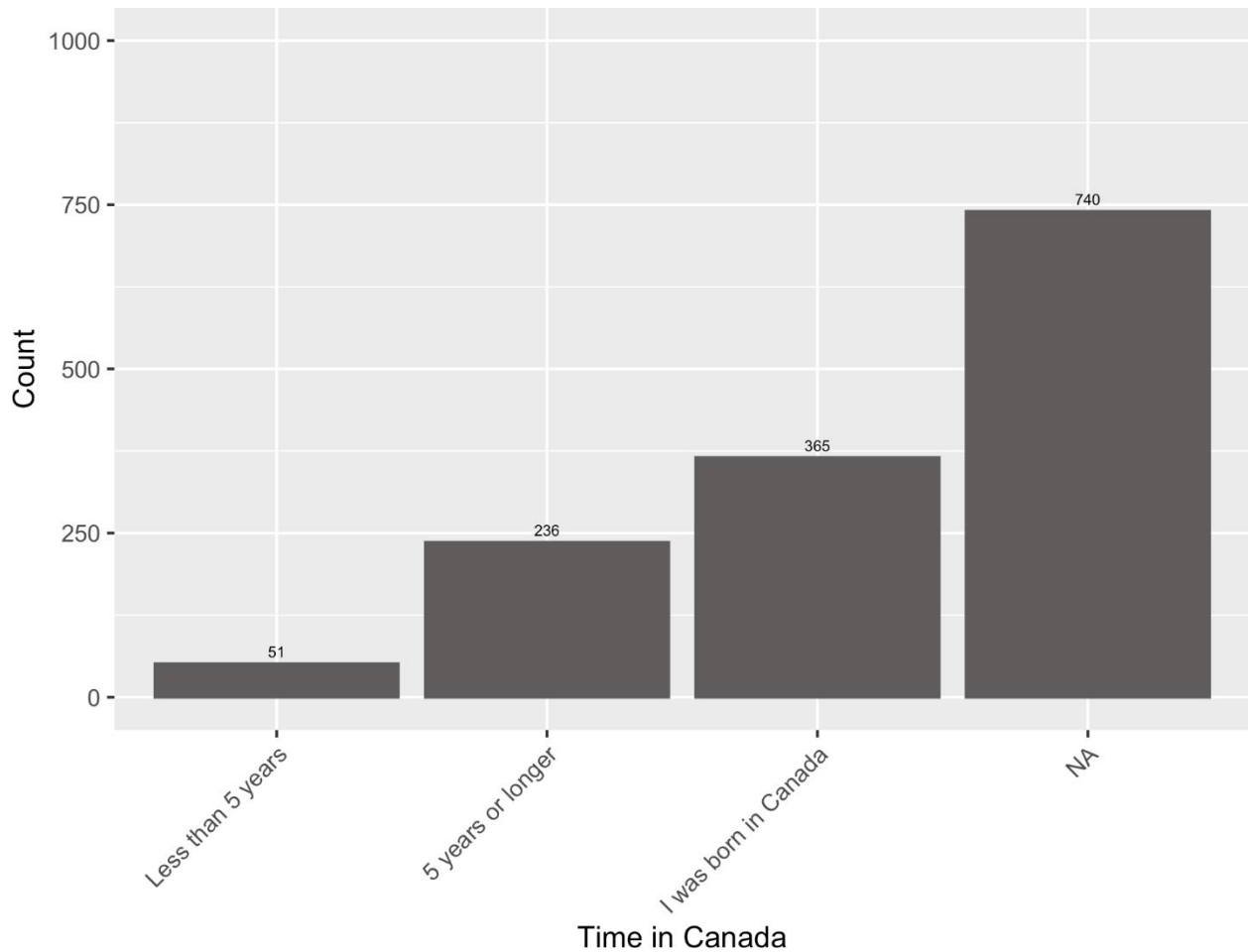


Figure 4: Number of Winter 2025 instructors by self-reported time in Canada. Note: NA represents Winter 2025 instructors who did not complete the Equity Survey or preferred not to answer the time in Canada question.

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 Indigenous 100% were born in Canada, and of those who are white, three-quarters (75%) were born in Canada. This is in contrast to less than one-quarter of racialized respondents (12%) who were born in Canada. For those respondents born outside of Canada, over two-thirds of those who are also racialized (71%) report being in Canada 5+ years, compared to less than one-quarter of those who are white (21%) who report being in Canada 5+ years. Notably, 41% of those who indicated they had been in Canada 5+ years selected ‘prefer not to answer’ when asked their self-reported racial identity, while 5% of those who selected ‘prefer not to answer’ indicated they had been in Canada less than 5 years.

An Analysis of Winter 2025 Student Course Perceptions Survey Responses

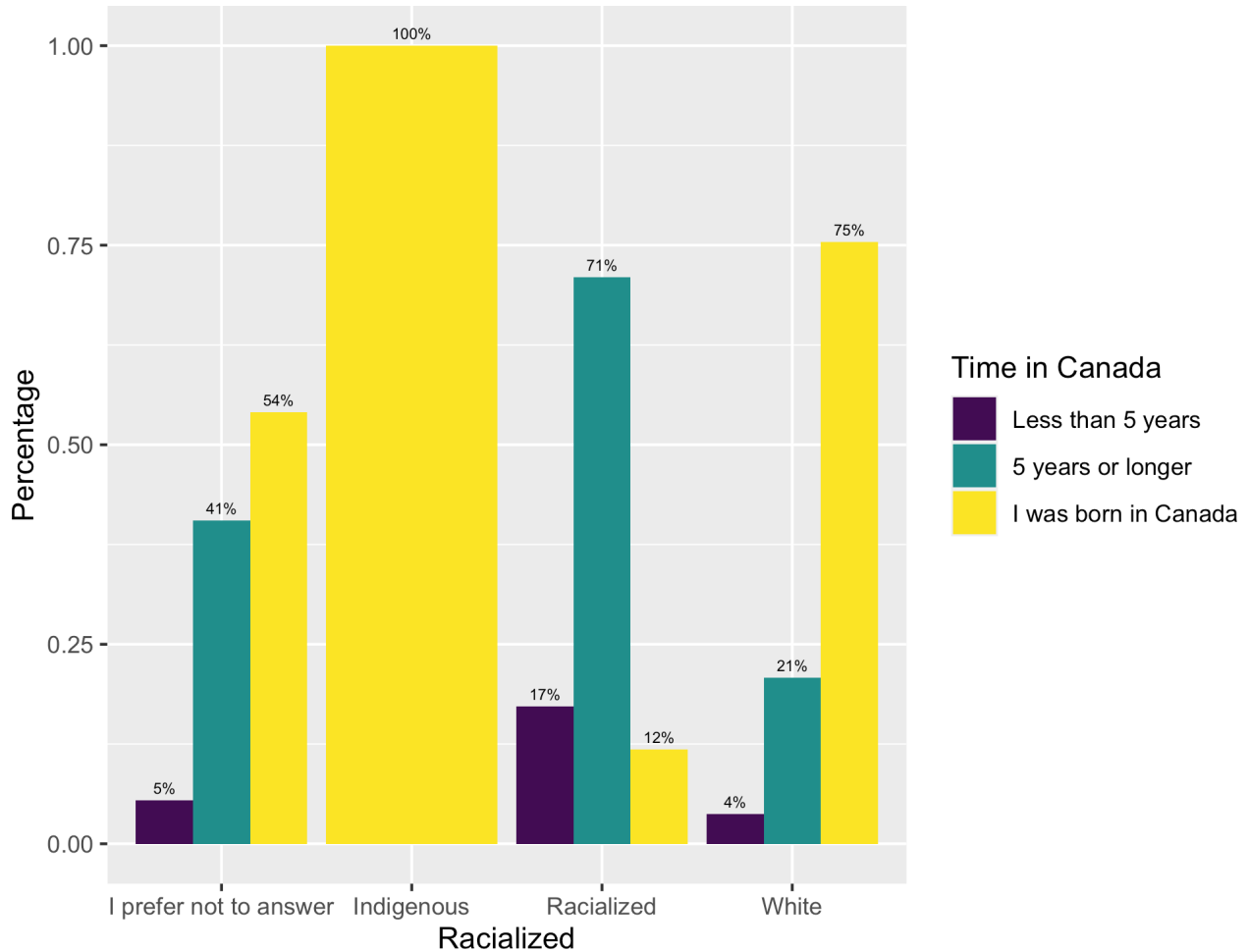


Figure 5: Self-reported Number of white, racialized & Indigenous Winter 2025 instructors by time spent in Canada.

Instructor-level Variable: Gender

Instructor Gender

Previous reports have shared descriptive data for Instructor sex based on HR data (the only data source available for that time period). For the 2025 analysis in addition to HR data for instructor sex, we also report descriptive data from the Equity survey for self-reported instructor gender. Figure 6 shows 53% of the respondents selected ‘other’¹, 23% selected ‘man’, 20% selected ‘woman’ and the remaining 2% selected ‘prefer not to answer.’

¹ The “Other” category includes respondents who either selected a self-described gender identity (e.g., agender, gender non-conforming, non-binary, questioning, trans, Two-Spirit, or another identity not explicitly listed) or selected multiple gender categories for this survey question.

An Analysis of Winter 2025 Student Course Perceptions Survey Responses

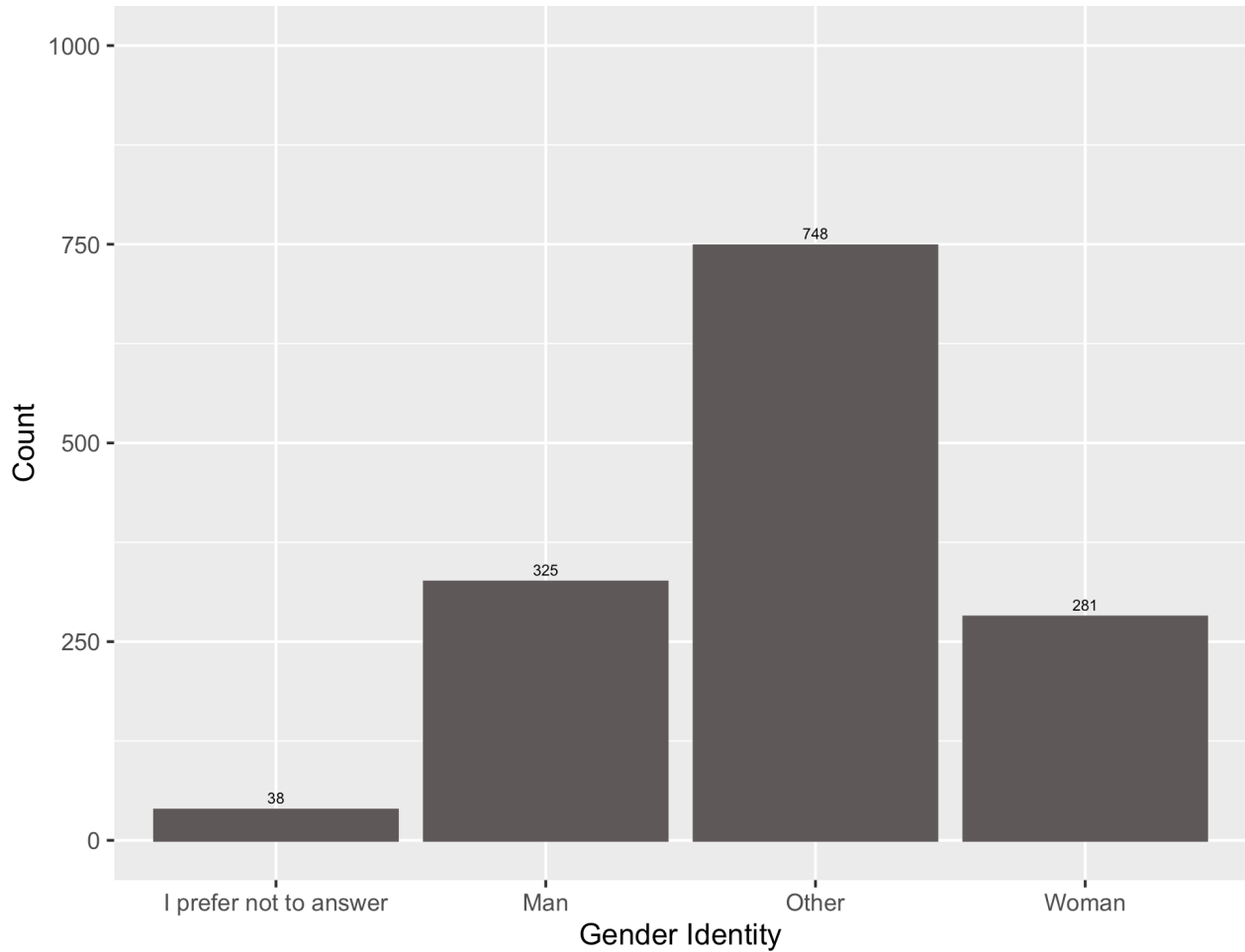


Figure 6: Self-reported Gender Identity for Instructors from 2025 Equity Survey

While self-reported gender identity provides insight into self-reported instructor gender identity, the distribution of responses, particularly the large proportion categorized as “Other,” which includes both self-reported identities and multiple selections, significantly limits further analysis for patterns related to self-reported gender identity.

To support analysis of representation in relation to institutional workforce patterns, as previous reports have relied on instructor sex at birth, we also examine HR data for 2025. This dataset provides consistent classification across the instructor population and allows for larger sample comparisons across faculties and appointment types. For this reason, the following sections focus on HR sex-at-birth data.

Instructor Sex

Turning to HR data which provides data for instructor ‘sex at birth’ Figure 7 shows 60% of 2025 instructors were male, while the remaining 40% were female.

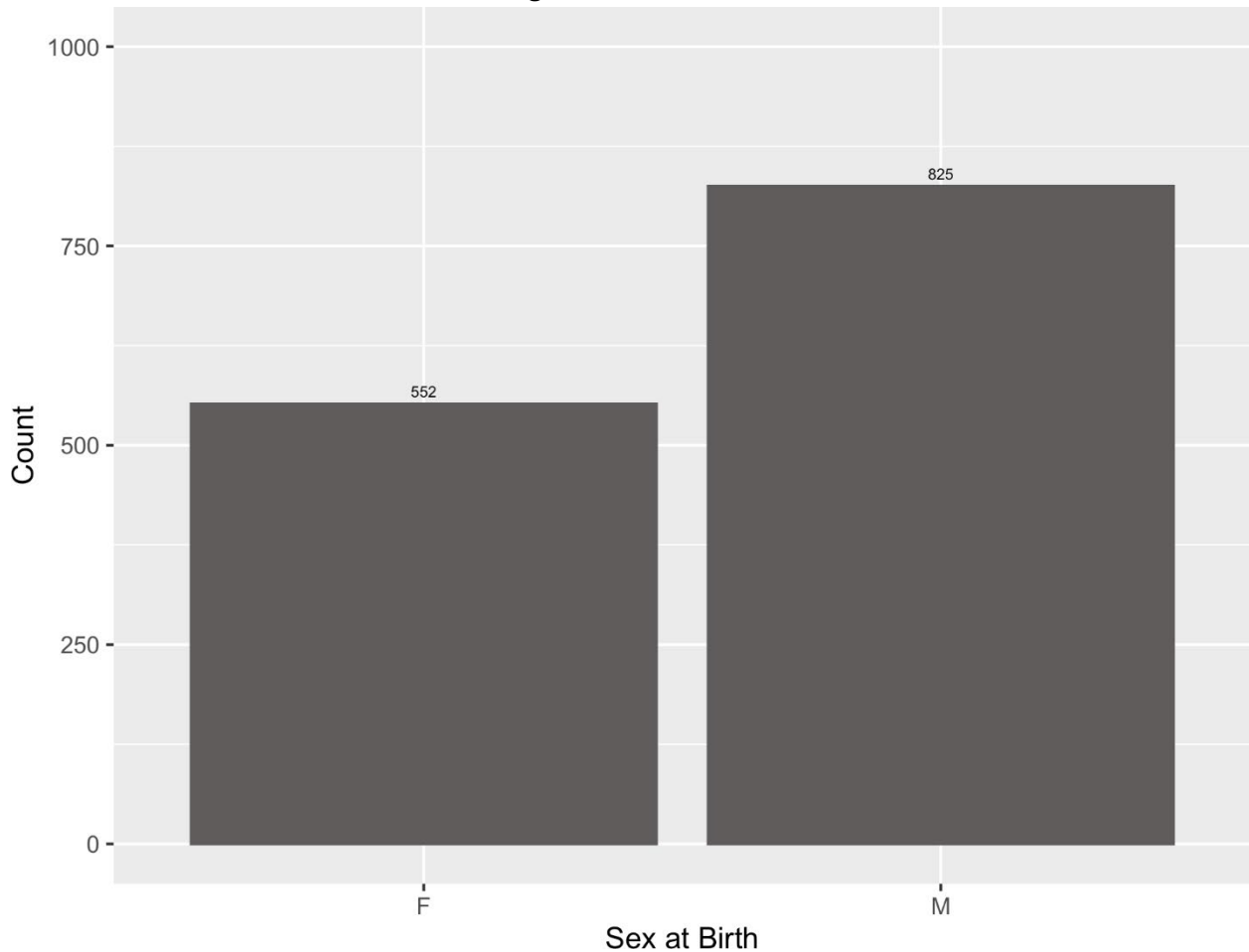


Figure 7: Number of Winter 2025 Instructors by Sex

Instructor Sex by Faculty

Figure 8 presents the distribution of instructors by faculty based on HR-reported sex at birth. Across most faculties, instructors recorded as male outnumber those recorded as female, although the magnitude of this difference varies by faculty.

The largest disparities are observed in Engineering and Mathematics. In Engineering, the number of instructors recorded as male (204) is nearly three times that of those recorded as female (76). A similar pattern is observed in Mathematics (199 male; 65 female). Science also shows a substantial difference (106 male; 71 female).

In contrast, Arts shows near parity between groups (219 male; 211 female), and Health is the only faculty where instructors recorded as female slightly outnumber those recorded as male (44 female; 40 male). Environment also shows a relatively smaller gap (52 male; 40 female).

An Analysis of Winter 2025 Student Course Perceptions Survey Responses

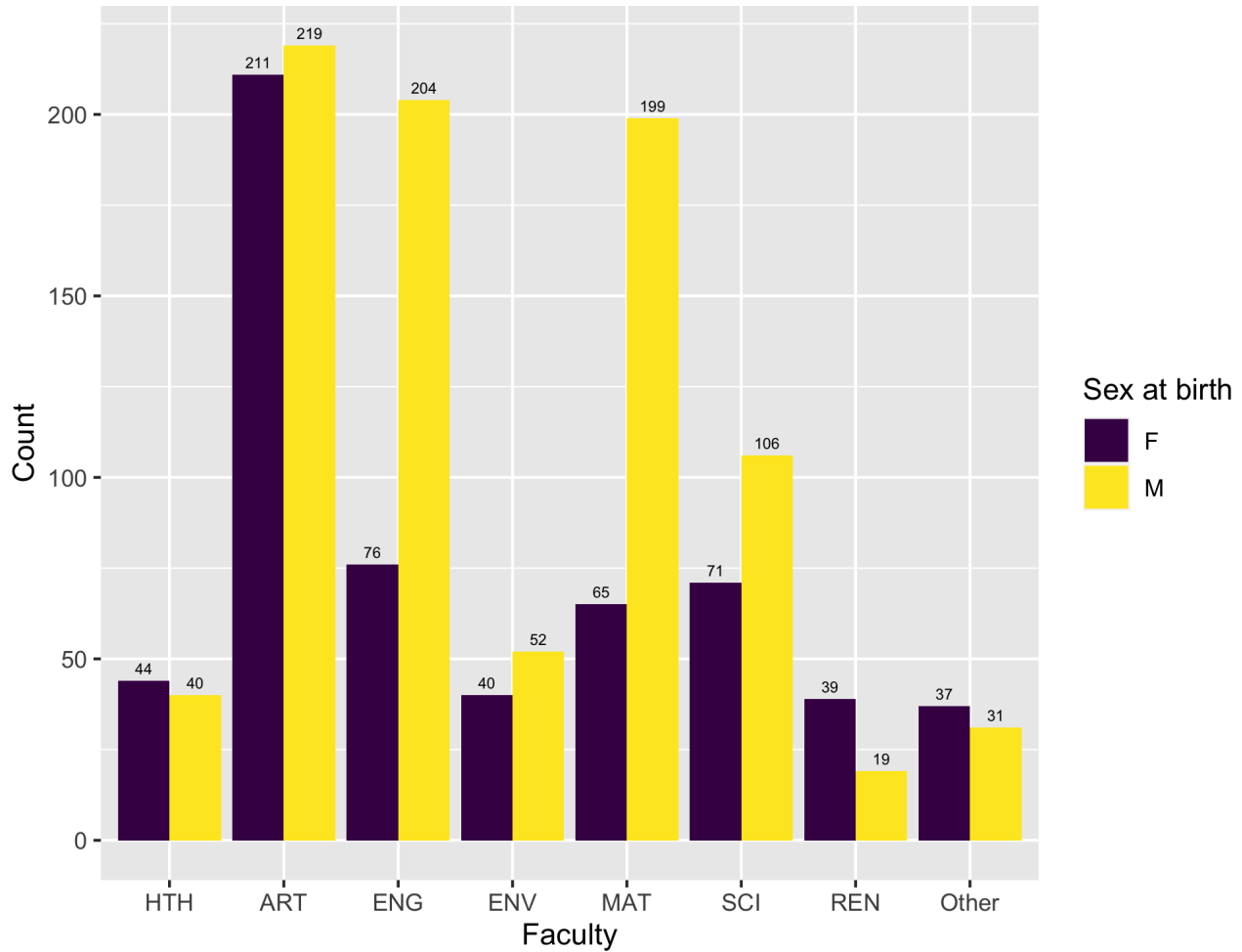


Figure 8: Number of Winter 2025 instructors by Faculty and Sex.

Instructor Sex by Appointment Type

Figure 9 further breaks down our sample characteristics by Winter 2025 instructor appointment type and sex at birth. In general, male instructors outnumber female instructors across all instructor types. Over four times more Winter 2025 Professors were male (181 males, 44 females) and two times more Winter 2025 Associate Professors were male (162 males, 85 females). The makeup of Winter 2025 instructors in remaining instructor types was less drastically imbalanced: Assistant Professors (70 males, 55 females), and Sessional (144 males, 118 females).

An Analysis of Winter 2025 Student Course Perceptions Survey Responses

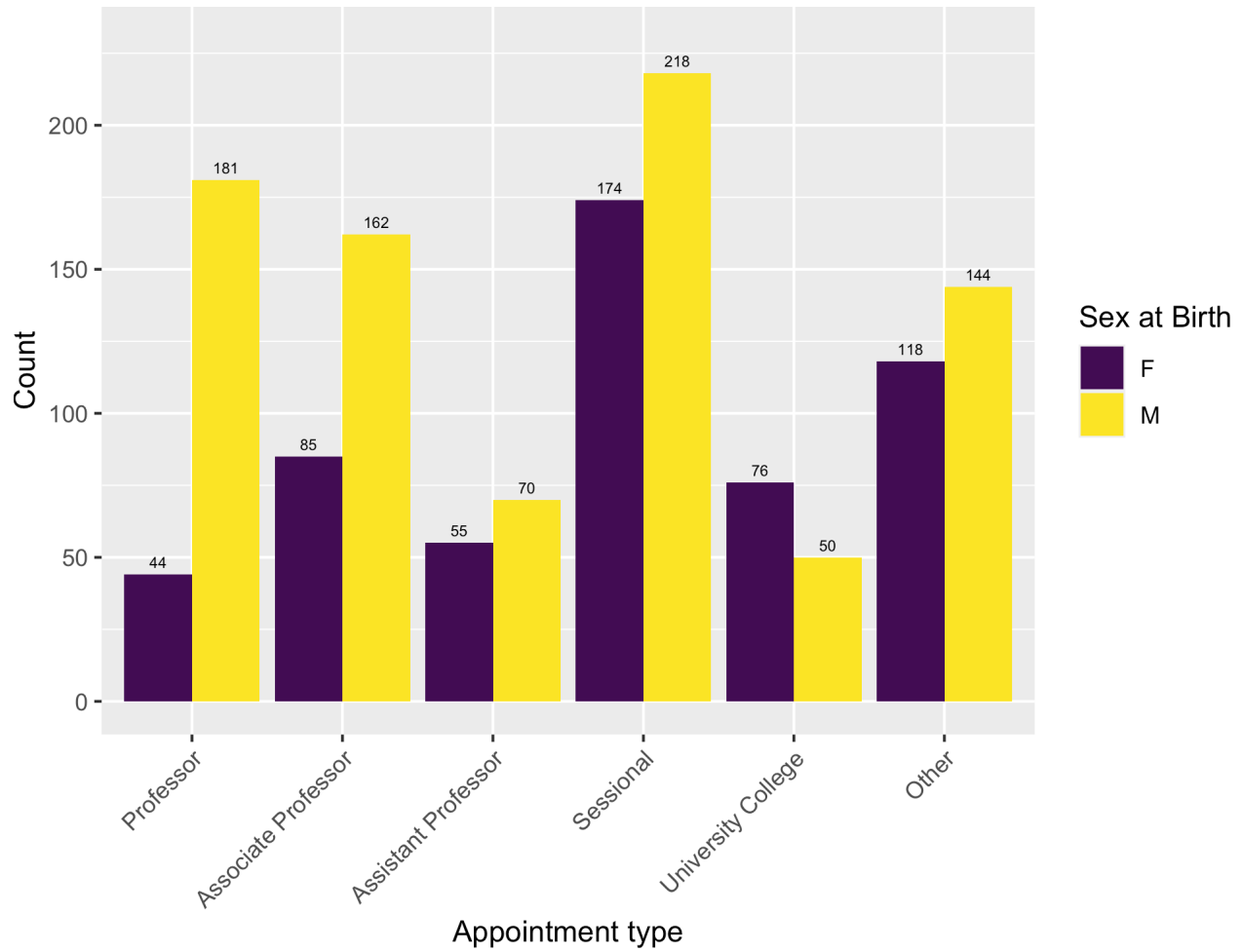


Figure 9: Number of Winter 2025 instructors by appointment type and sex at birth.

Course-level Variable: Class Size

Figure 10 outlines the number of Winter 2025 courses included in this study by class size. The bulk of courses (2,254) had 100 or fewer students; only a small number (58) had 201+ students enrolled.

An Analysis of Winter 2025 Student Course Perceptions Survey Responses

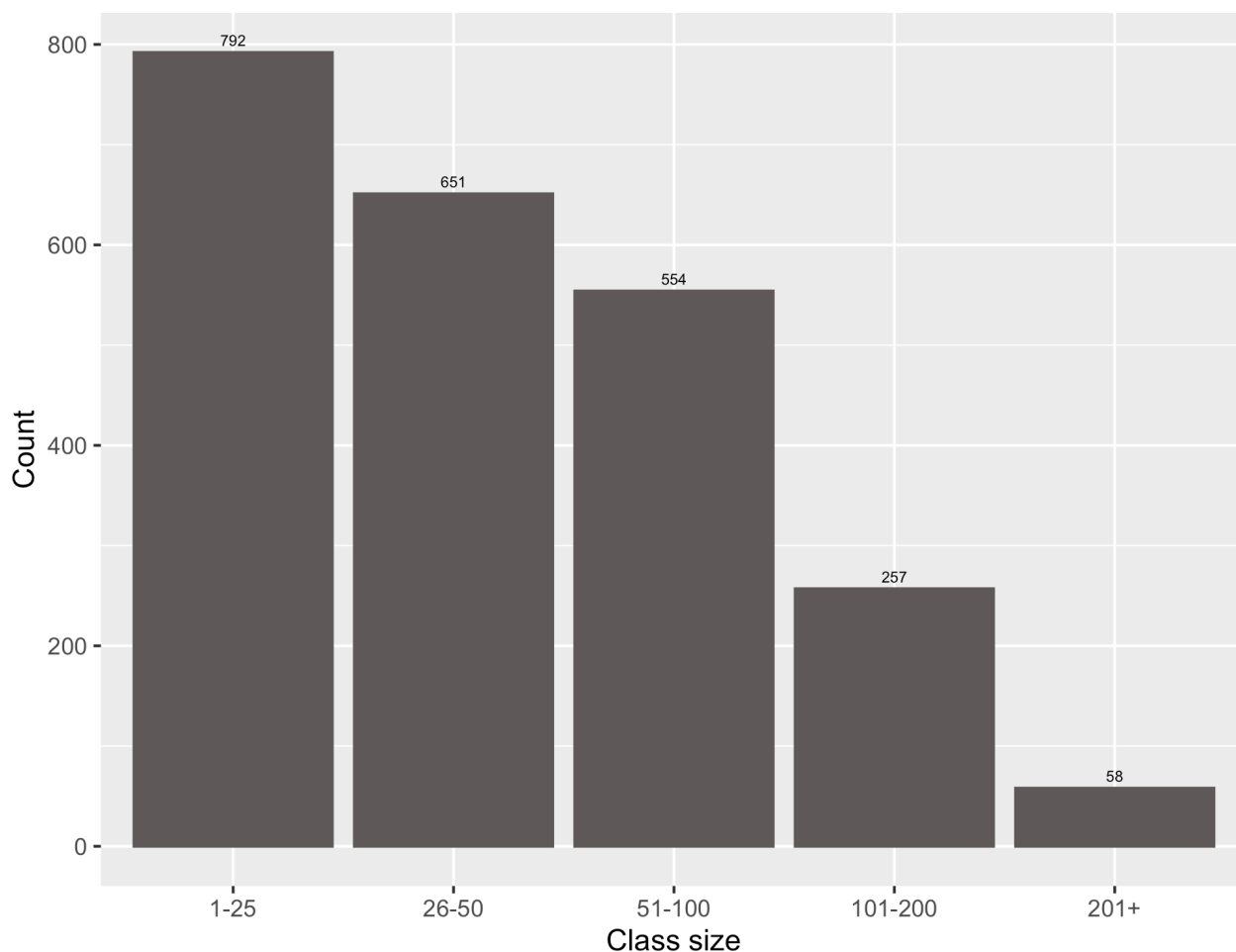


Figure 10: Number of Winter 2025 courses by class size.

Course-level Variable: Course Type by Faculty

The majority of Winter 2025 courses included in this study were held in person, as reflected in Figure 11. The Faculty of Arts had the highest number of online course offerings (106), but still held just over six times that number of courses in person (653). Engineering held the second highest number of in-person courses (454), with only 18 courses offered online and Science held the least online courses (12) in contrast to 269 in person.

An Analysis of Winter 2025 Student Course Perceptions Survey Responses

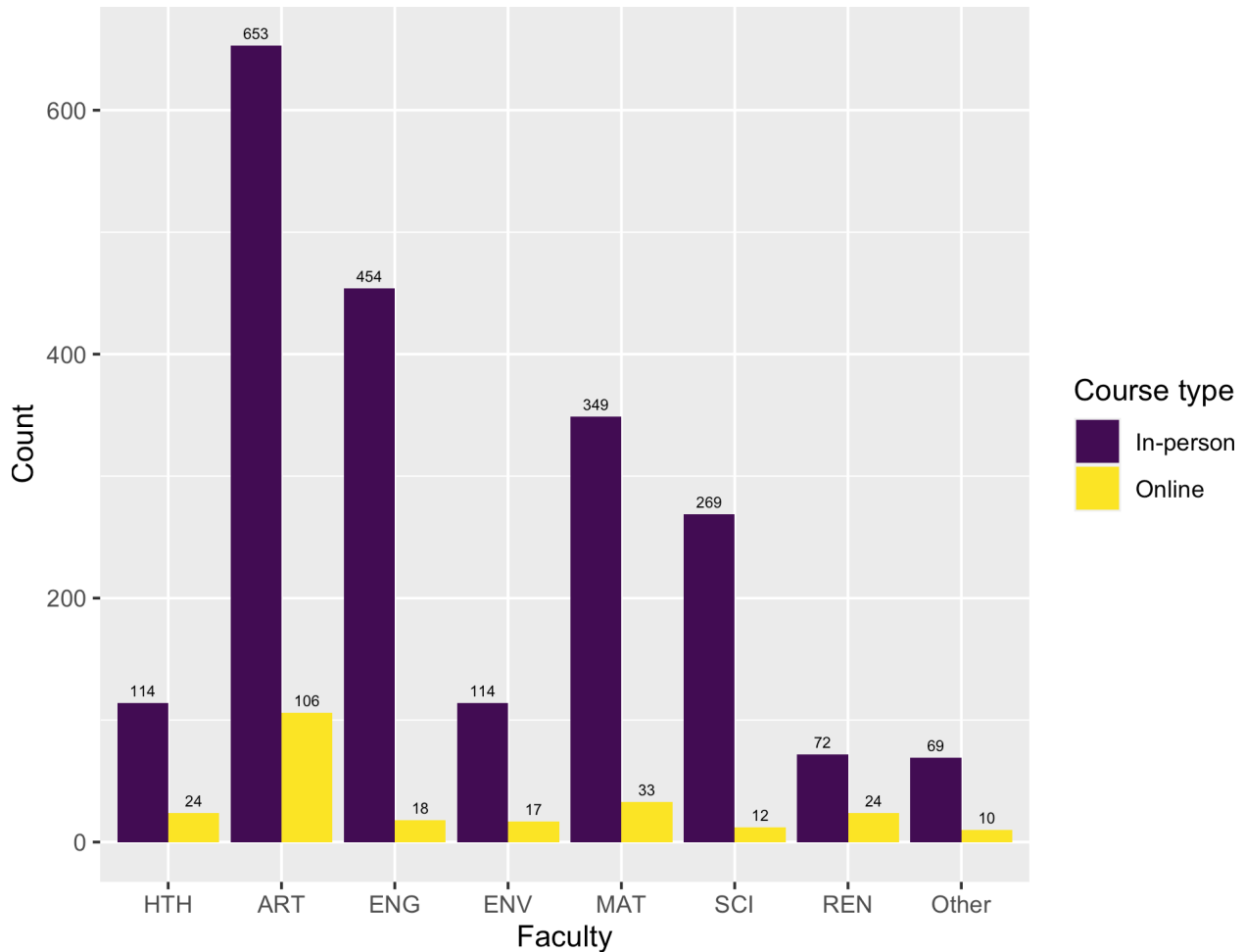


Figure 11: Number of Winter 2025 course sections by faculty and course type.

SCP Responses by Response Item

Figure 12 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 72% and 84% of the time across all six items. Conversely, respondents selected ‘Disagree’ and ‘Strongly Disagree’ between 5% and 11% 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 9% and 14% of the time across all six items.

An Analysis of Winter 2025 Student Course Perceptions Survey Responses

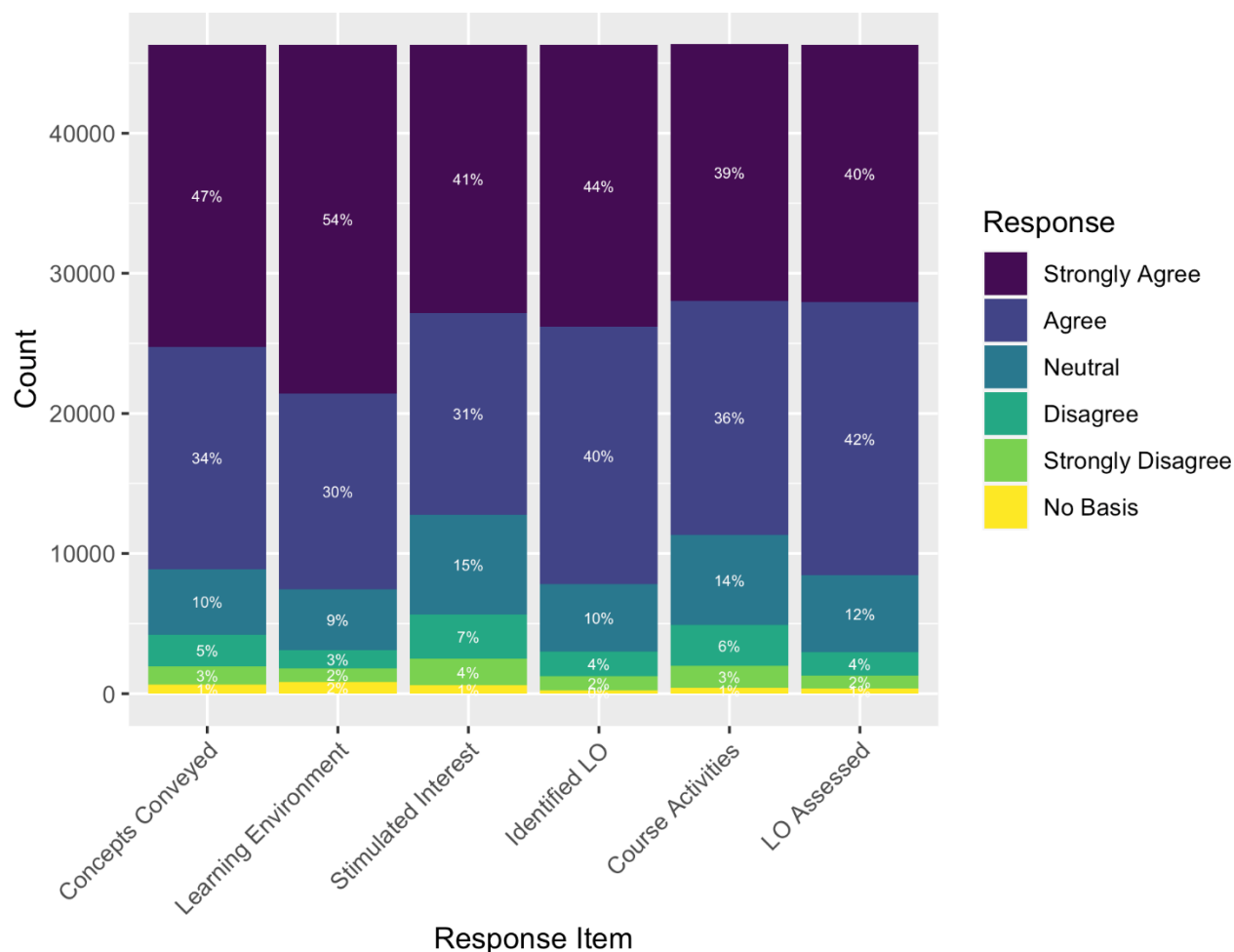


Figure 12: Proportion of Winter 2025 responses per response item.

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 13 displays the mean rating for each of the six response items accounting for class size. Reiterating what we observed in Figure 12, 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.

The observed pattern is quite clear: 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 3.9 in 200+ courses to close to 4.6 in courses with 25 or fewer students). Conversely, *Stimulated_Interest* (ranging from 3.7 for 200+ students to 4.3 for 1-25 students) and *Course_Activities* (ranging from 3.9 for 200+ students to 4.3 for 1-25 students) receive the

lowest mean rating across all class sizes.

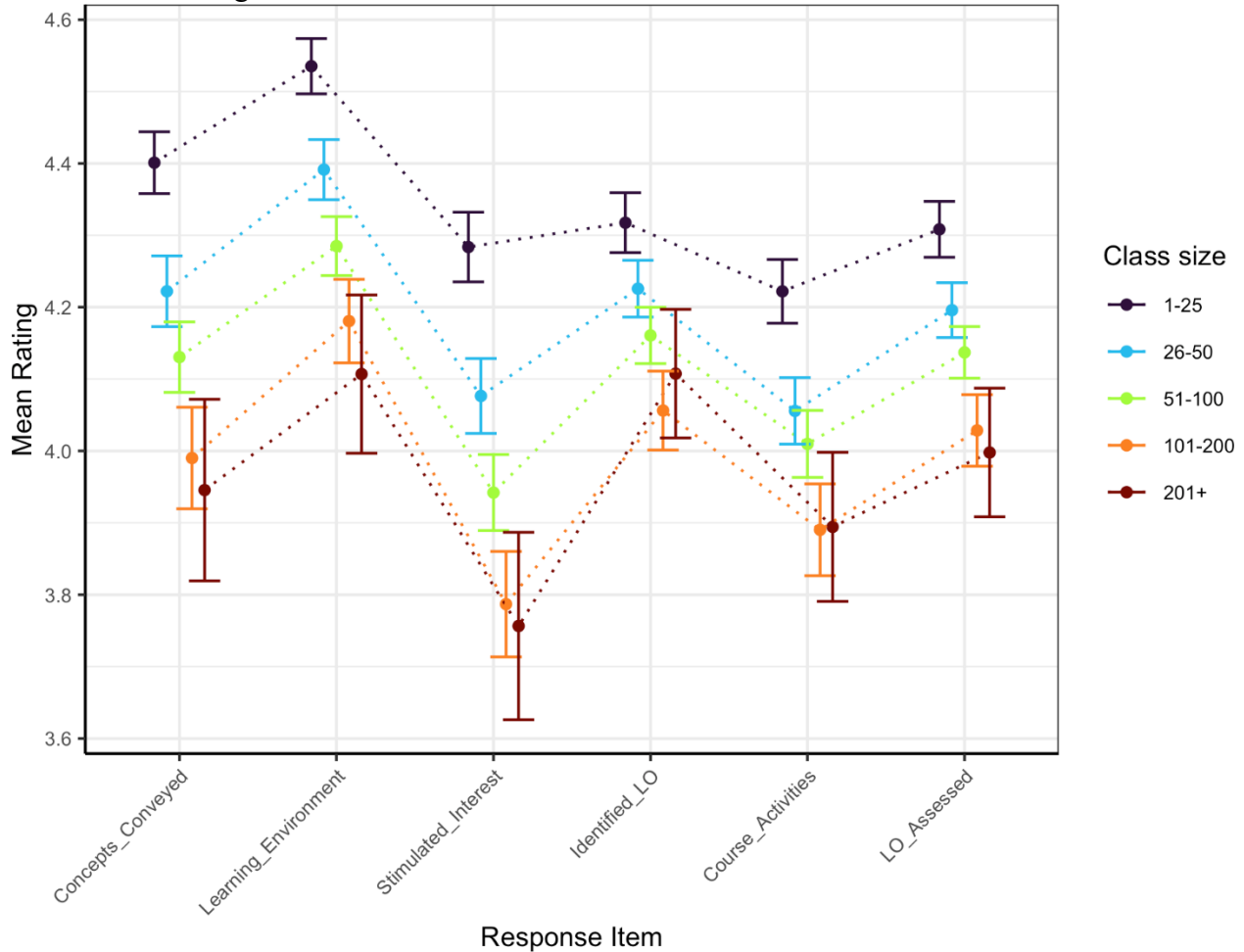


Figure 13: Mean Winter 2025 rating by response item and class size. Error bars correspond to 95% confidence intervals.

Mean Ratings by Response Item and Faculty of Course Offering

Figure 14 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 4.4 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 about 4.0 and in Arts was 4.1, which translates to a negligible difference – just 0.1 on the 5-point scale.

An Analysis of Winter 2025 Student Course Perceptions Survey Responses

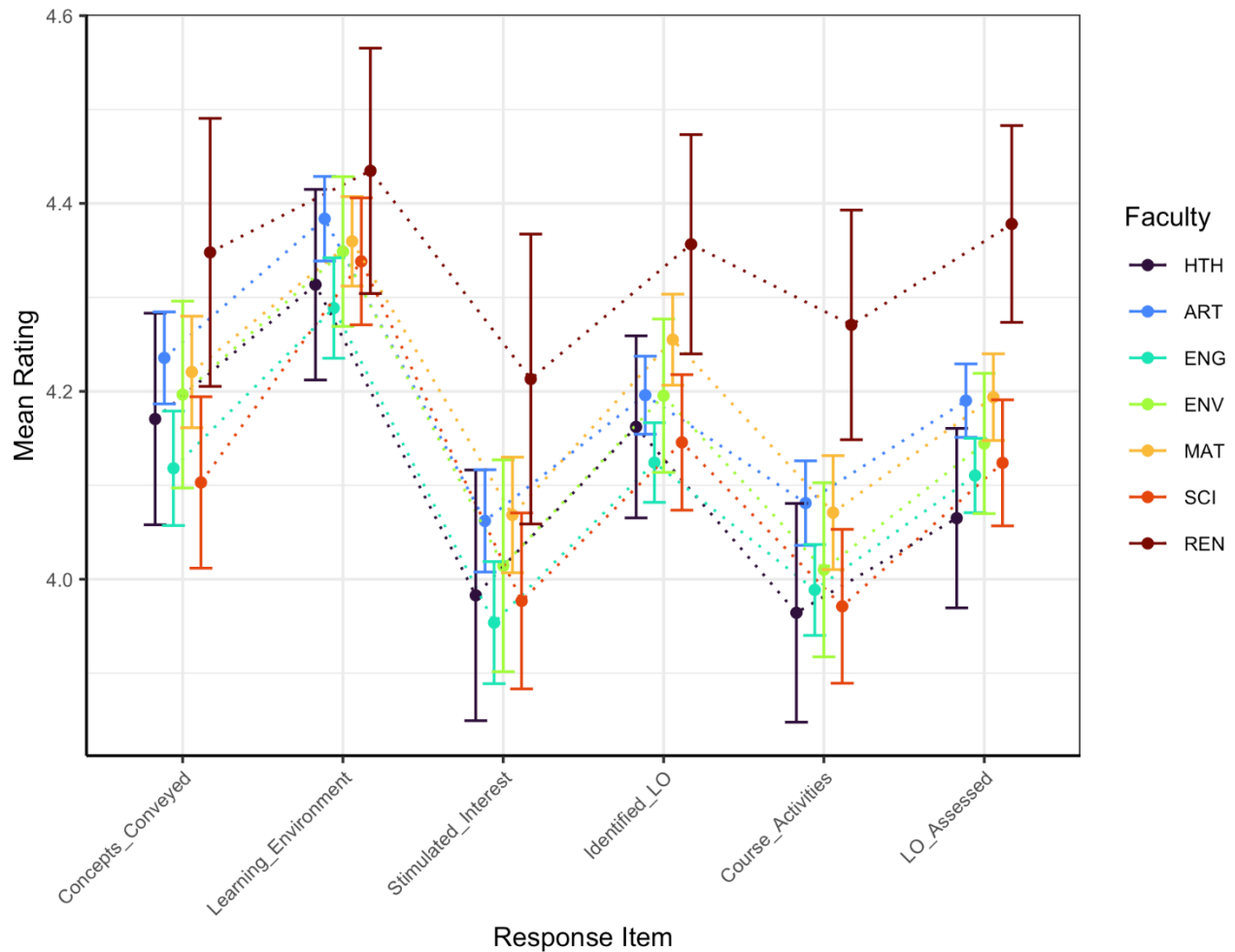


Figure 14: Mean Winter 2025 rating by response item and Faculty of course offering. Error bars correspond to 95% confidence intervals.

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 2025 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, (as evidenced by the error bar crossing or touching the zero line). Across nearly all cases, the non-significant difference is extremely small (as evidenced by the data points falling close to the zero line).

An Analysis of Winter 2025 Student Course Perceptions Survey Responses

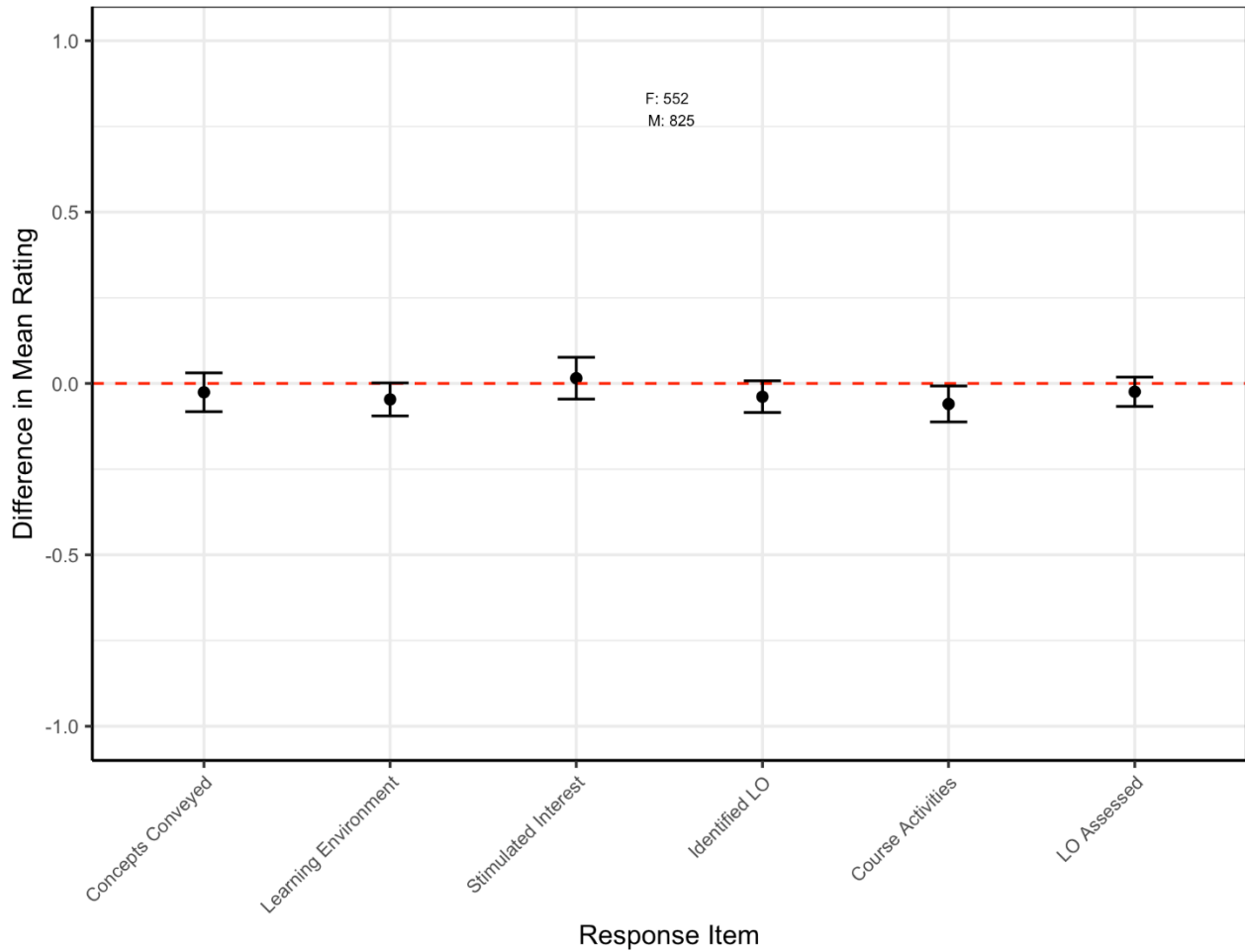


Figure 15: Difference in mean Winter 2025 ratings for male and female instructors per response item. Error bars correspond to 95% confidence intervals.

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 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 2025 Student Course Perceptions Survey Responses

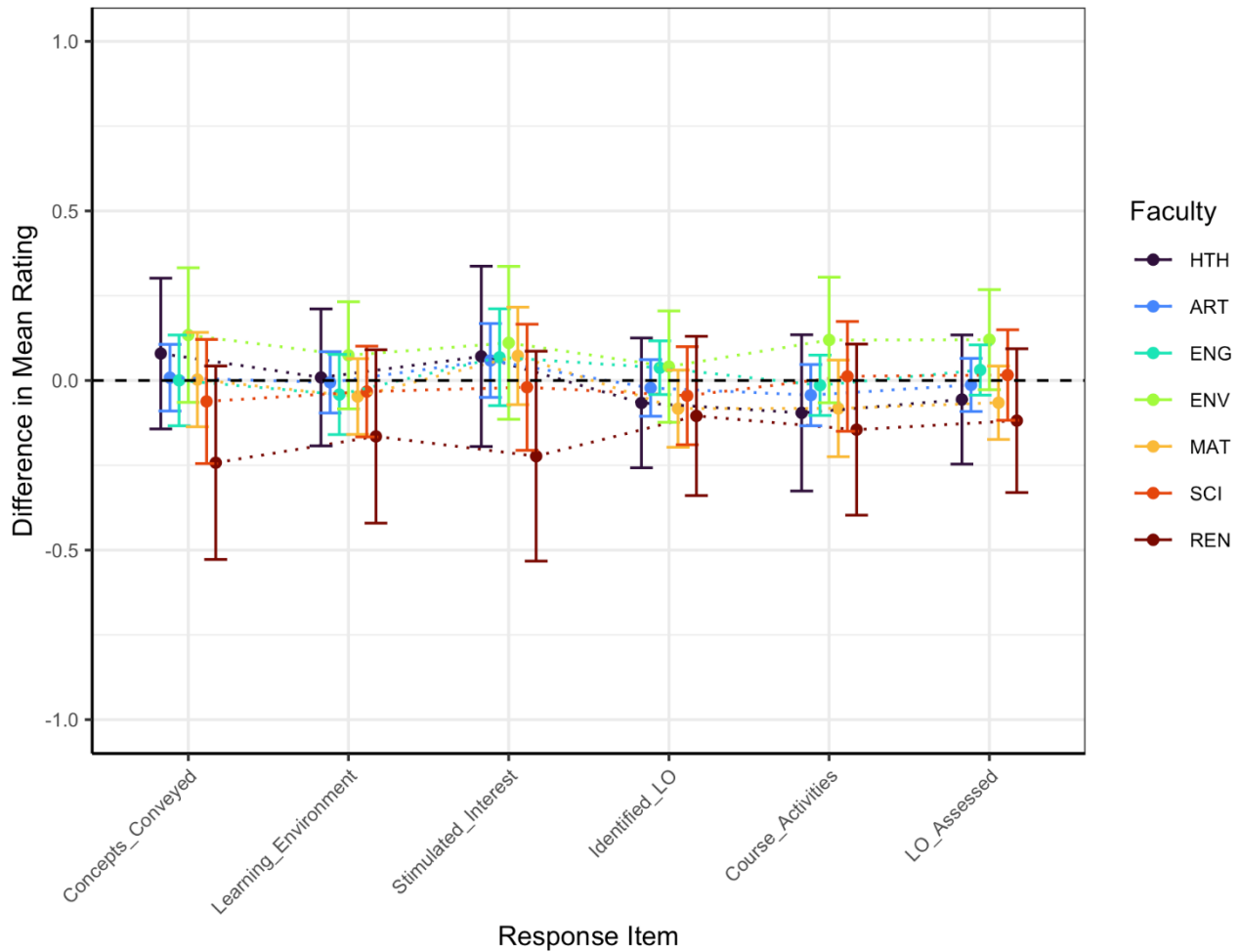


Figure 16: Difference in mean Winter 2025 ratings for male and female instructors by response item and Faculty of course offering. Error bars correspond to 95% confidence intervals.

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

Figure 17 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

An Analysis of Winter 2025 Student Course Perceptions Survey Responses

in most cases, the non-significant difference is very small (as evidenced by the data points falling close to the zero line).

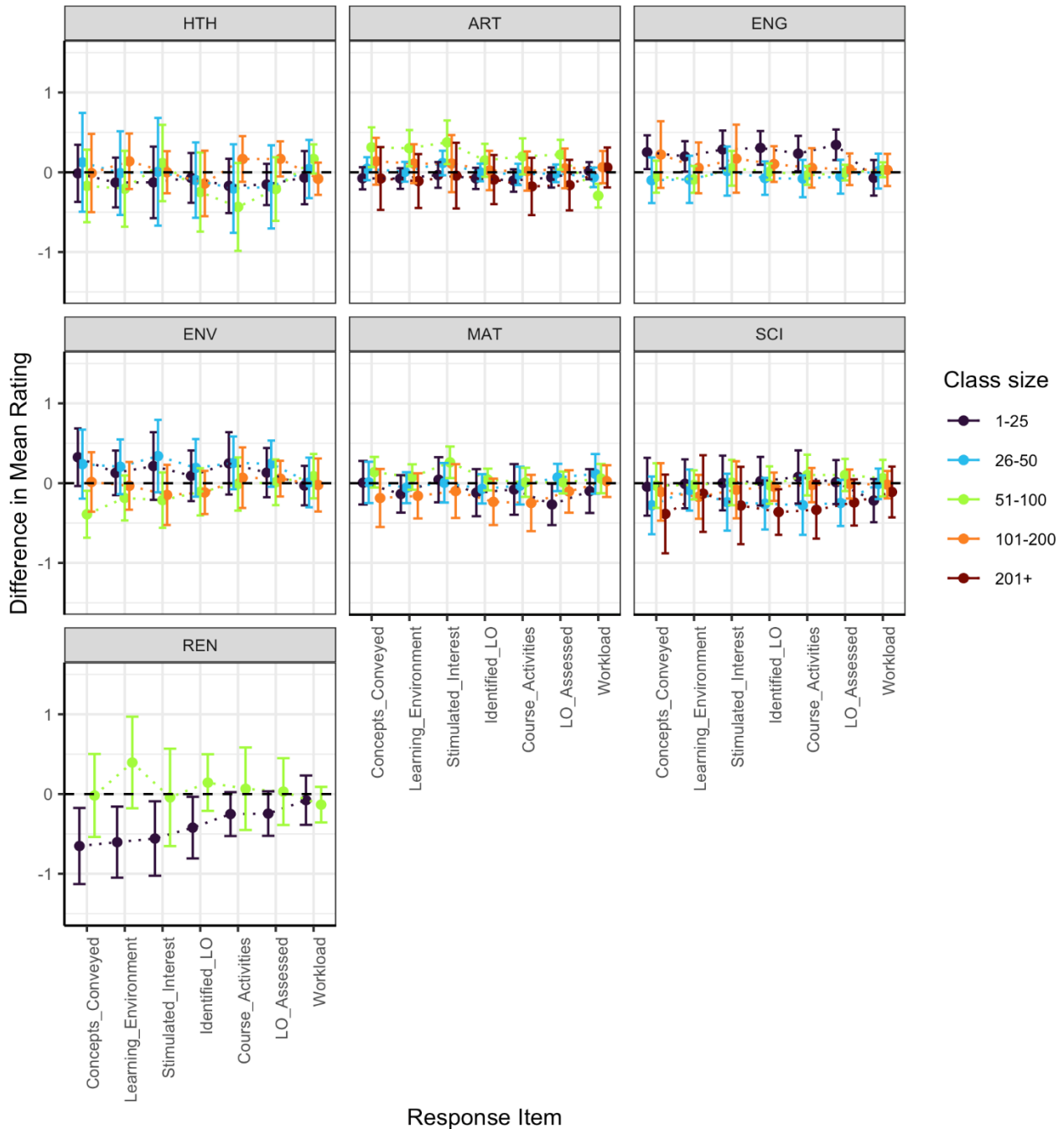


Figure 17: Difference in mean Winter 2025 ratings between male and female instructors per response item by class size and Faculty. Error bars correspond to 95% confidence intervals.

A few exceptions to this general pattern are observed. In Arts, for classes of 51–100 students, statistically significant differences appear for Concepts_Conveyed, Stimulated_Interest, Identified_LO, and Workload. For the first three items, error bars lie above the zero line, indicating that male instructors score slightly higher than female instructors (by approximately

0.5 points or less). In contrast, for Workload, the error bar falls below zero, indicating a similarly small advantage for female instructors.

In Engineering, statistically significant differences are observed in smaller classes (1–25 students) for Concepts_Conveyed, Stimulated_Interest, Identified_LO, and LO_Assessed. In each case, error

bars are above zero, indicating that male instructors again receive modestly higher ratings (≤ 0.5 points).

In Environment, only one significant difference is observed: in classes of 51–100 students, female instructors score just under 0.5 points higher on Concepts_Conveyed.

In Mathematics, a significant difference appears in classes of 51–100 for Stimulated_Interest, with male instructors scoring slightly lower (≤ 0.5 points) than female instructors.

Finally, in Science, for large classes (201+ students), a statistically significant difference is observed for Identified_LO, with male instructors scoring modestly higher (less than 0.5 points).

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

We also examined differences in mean ratings by course type (in-person vs. online) across Faculties (Figure 18). Overall, most differences are small in magnitude, and many are not statistically significant, particularly for in-person courses, where estimates tend to cluster tightly around zero.

A more variable pattern emerges for online courses. In several Faculties, observed differences are larger and, in some cases, statistically significant. These differences are most pronounced in Environment and Science. In Environment, online courses show consistently higher ratings for male instructors across most items, with differences approaching up to approximately one point on some measures (e.g., Concepts_Conveyed and Stimulated_Interest). In Science, the pattern is reversed: online courses show lower ratings for male instructors across most items, with differences in the range of roughly 0.5 to 1 point.

In other Faculties (e.g., Health, Arts, Mathematics, and Renison), differences in online courses are more modest and less consistent across items, with confidence intervals often overlapping zero. Engineering shows minimal variation across both instructional modes.

Importantly, even where statistically significant differences are observed, these patterns are not uniform across Faculties or survey items, and in some cases differ in direction. This variability suggests that the relationship between instructor gender and student ratings may be context-dependent and influenced by factors that vary by discipline and mode of delivery. As such, these findings should be interpreted cautiously and not as evidence of a consistent or systemic pattern across the institution.

An Analysis of Winter 2025 Student Course Perceptions Survey Responses

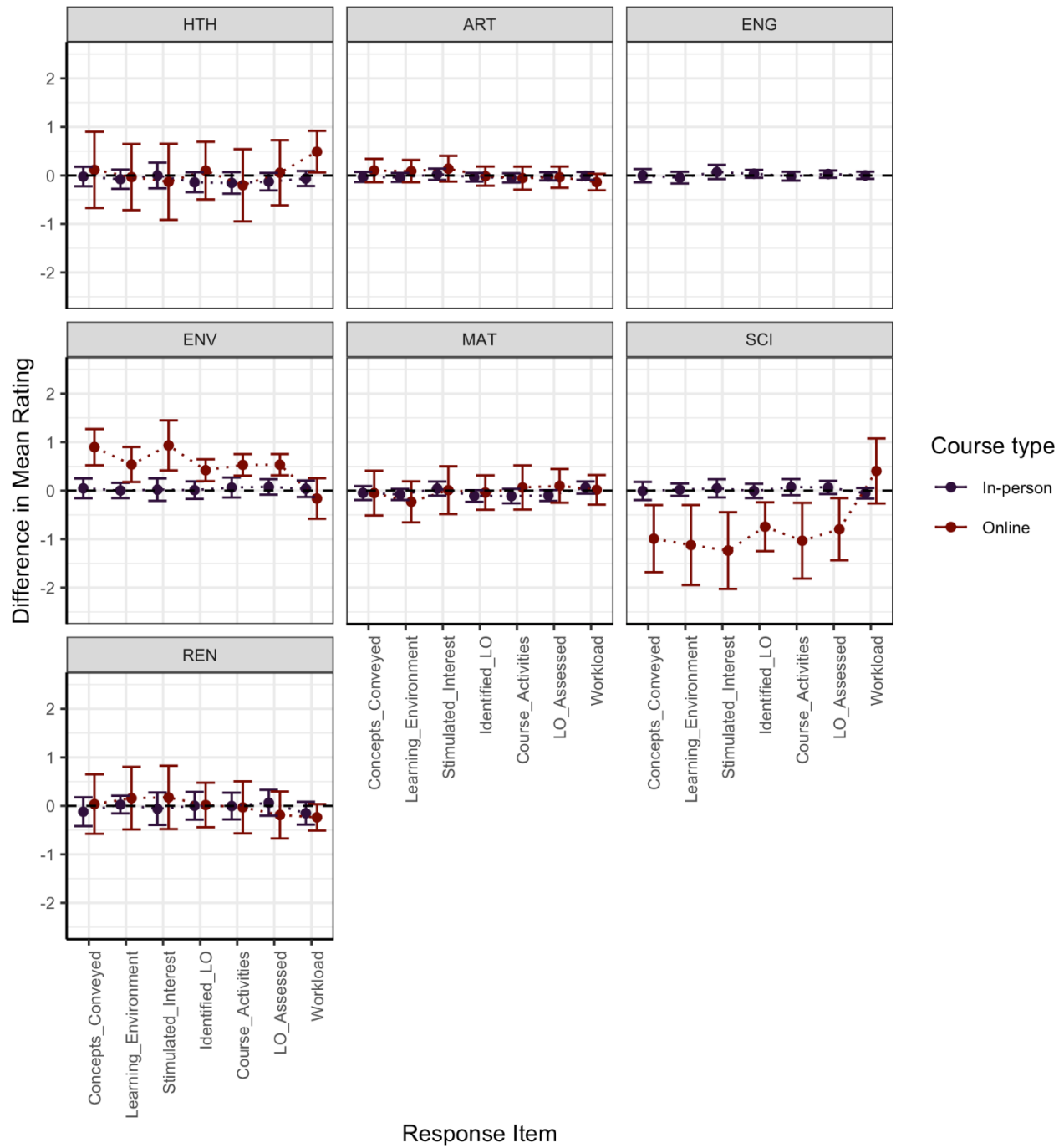


Figure 18: Difference in mean Winter 2025 ratings between male and female instructors per response item by course type and Faculty. Error bars correspond to 95% confidence intervals

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

Figure 19 shows differences in mean ratings between female and male instructors after accounting for instructor appointment type and class size. Overall, the results are remarkably consistent: differences are almost always statistically non-significant and cluster tightly around zero across all survey items, regardless of course size or appointment type.

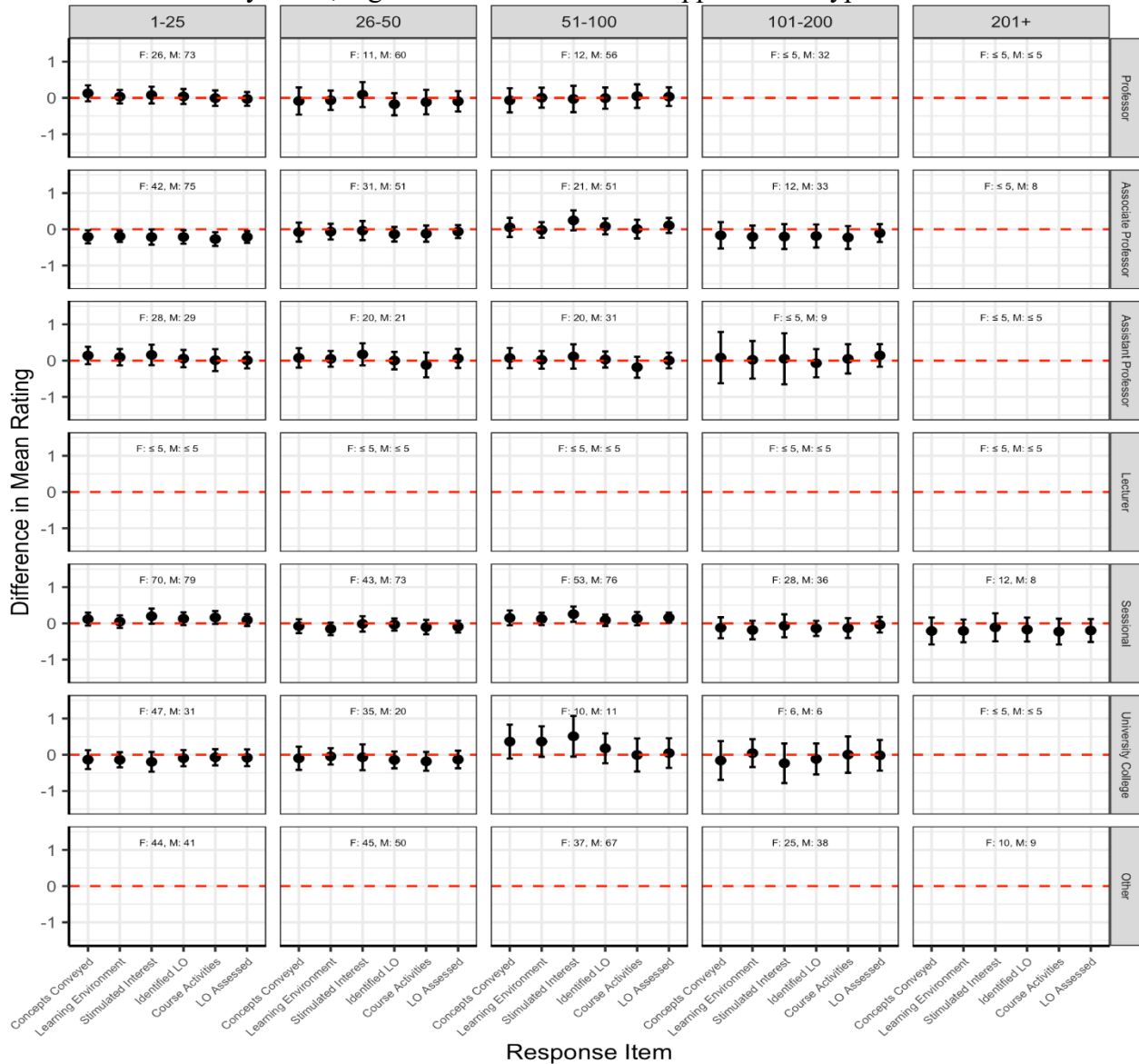


Figure 19: Difference in Winter 2025 mean rating for Male and Female Instructors by Class Size and Instructor Appointment Type. Error bars correspond to 95% confidence intervals.

Difference in MR for Male and Female Instructors by Class Size

Figure X examines differences in mean ratings between female and male instructors by class size across survey items.

A clear overall pattern emerges: differences are generally small and cluster close to zero, particularly in smaller and mid-sized classes. For classes with fewer than 50 students, estimates are tightly grouped around zero across all items, with confidence intervals consistently overlapping the zero line, indicating little evidence of meaningful or statistically reliable differences.

In mid-sized classes (51–100 students), a slightly different pattern appears. Across several items (e.g., Concepts_Conveyed and Stimulated_Interest), estimates tend to be modestly positive, suggesting somewhat higher ratings for male instructors, but the magnitude remains small (generally well below 0.25 points), and confidence intervals still overlap zero. This suggests that, while a directional pattern is visible, differences remain limited in substantive terms.

For larger classes (101–200 students), results remain close to zero and show no consistent directional trend across items. Differences vary slightly by measure but continue to be small and generally not statistically distinguishable from zero.

The largest deviations appear in the largest class size category (201+ students), where estimates are more consistently negative across items such as Learning_Environment, Course_Activities, and LO_Assessed. In this group, male instructors tend to receive lower ratings than female instructors, with differences approaching approximately -0.2 to -0.3 points. However, these estimates are accompanied by wider confidence intervals, suggesting greater variability and less precision in this category.

Taken together, the results indicate that class size does not introduce a clear or systematic pattern of gender-based differences in ratings. Where variation appears, it is modest, inconsistent across items, and often statistically uncertain. This reinforces the broader finding that differences are generally small and context-dependent rather than large or uniform..

An Analysis of Winter 2025 Student Course Perceptions Survey Responses

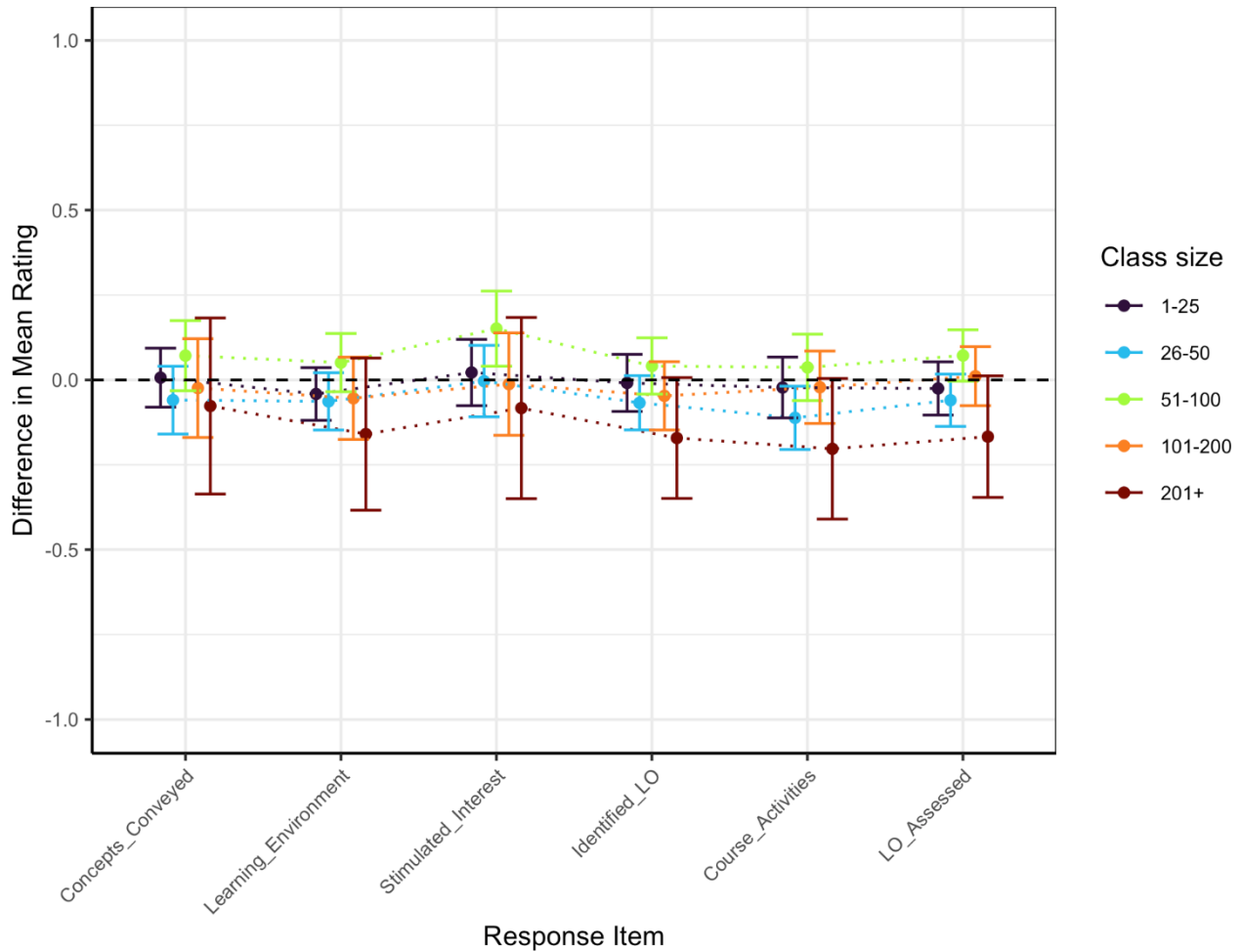


Figure 20: Difference in mean rating for Winter 2025 Male and Female Instructors by class size. Error bars correspond to 95% confidence intervals.

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

Figure 21 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 2025 Student Course Perceptions Survey Responses

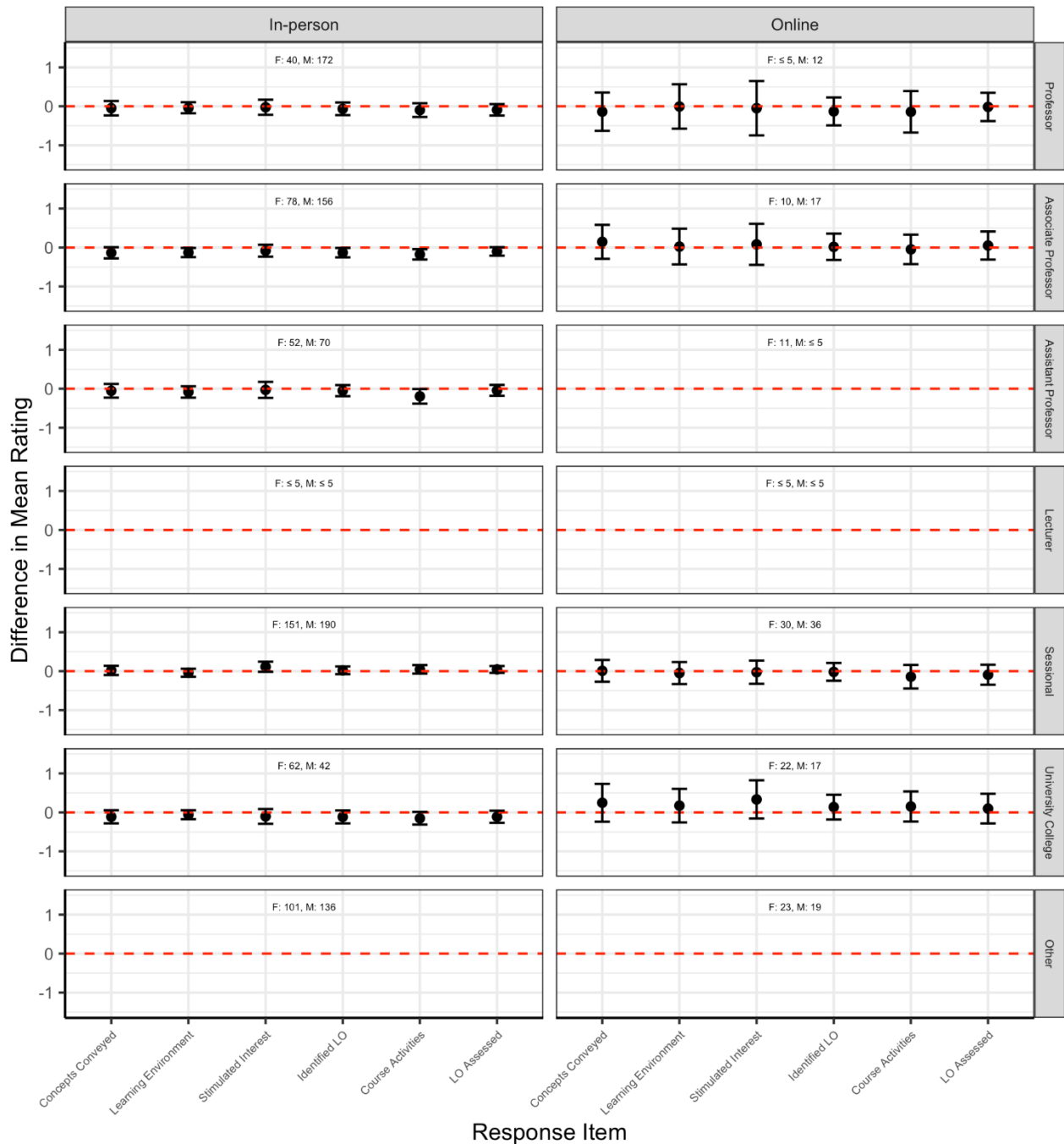


Figure 21: Difference in mean Winter 2025 ratings for male and female instructors by response item, instructor appointment type and course type. Error bars correspond to 95% confidence intervals.

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 22). The pattern is strikingly consistent: across all Faculties and appointment categories, differences are negligible. Estimates

An Analysis of Winter 2025 Student Course Perceptions Survey Responses

sit almost exactly on zero, and in every case, confidence intervals span the zero line, indicating no statistically reliable differences. In practical terms, once appointment type is taken into account, student ratings for male and female instructors are essentially indistinguishable.

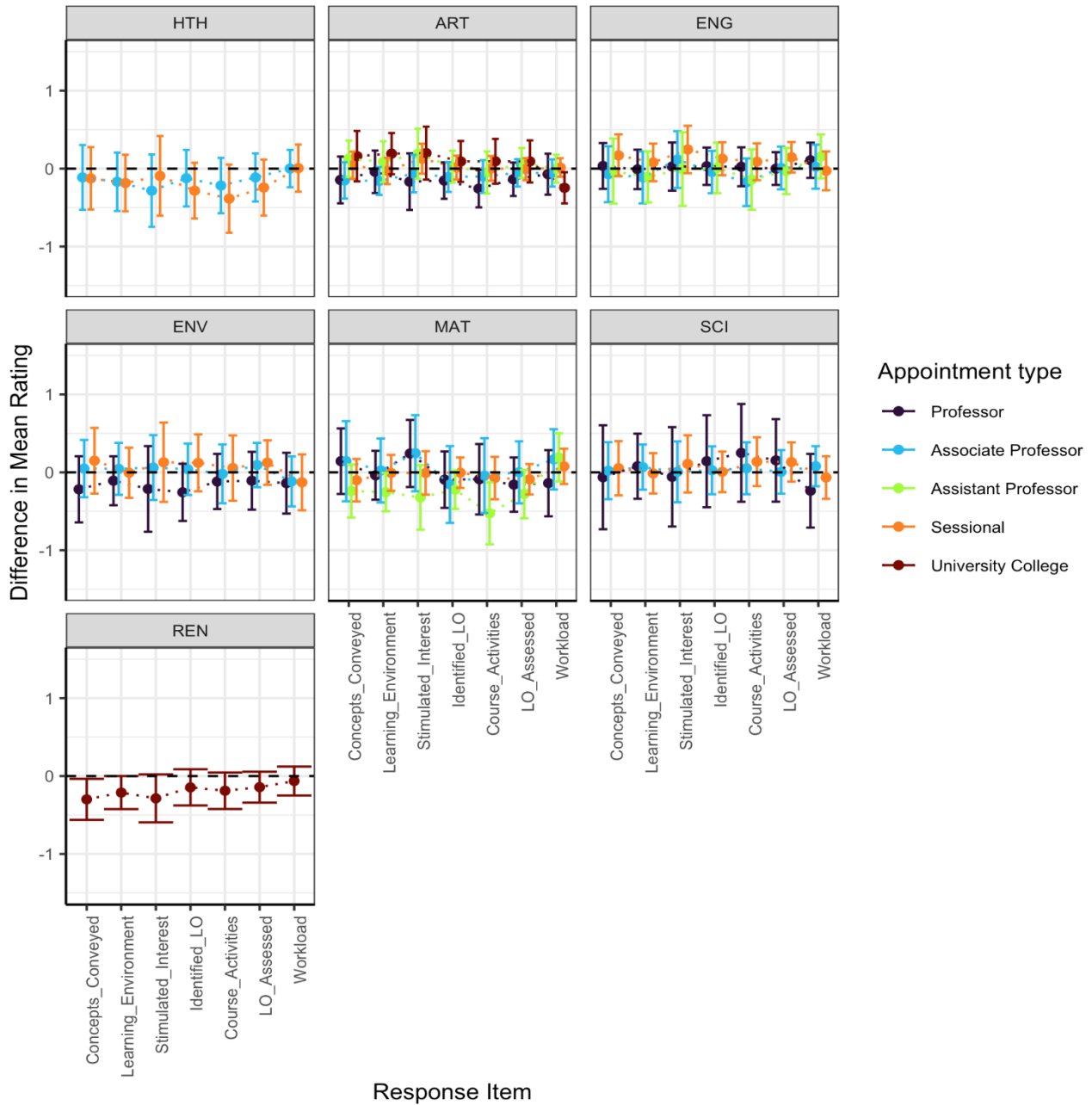


Figure 22: Difference in Winter 2025 mean rating between male and female instructors per response item by instructor appointment and Faculty. Error bars correspond to 95% confidence intervals.

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

Figure 23 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

An Analysis of Winter 2025 Student Course Perceptions Survey Responses

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).

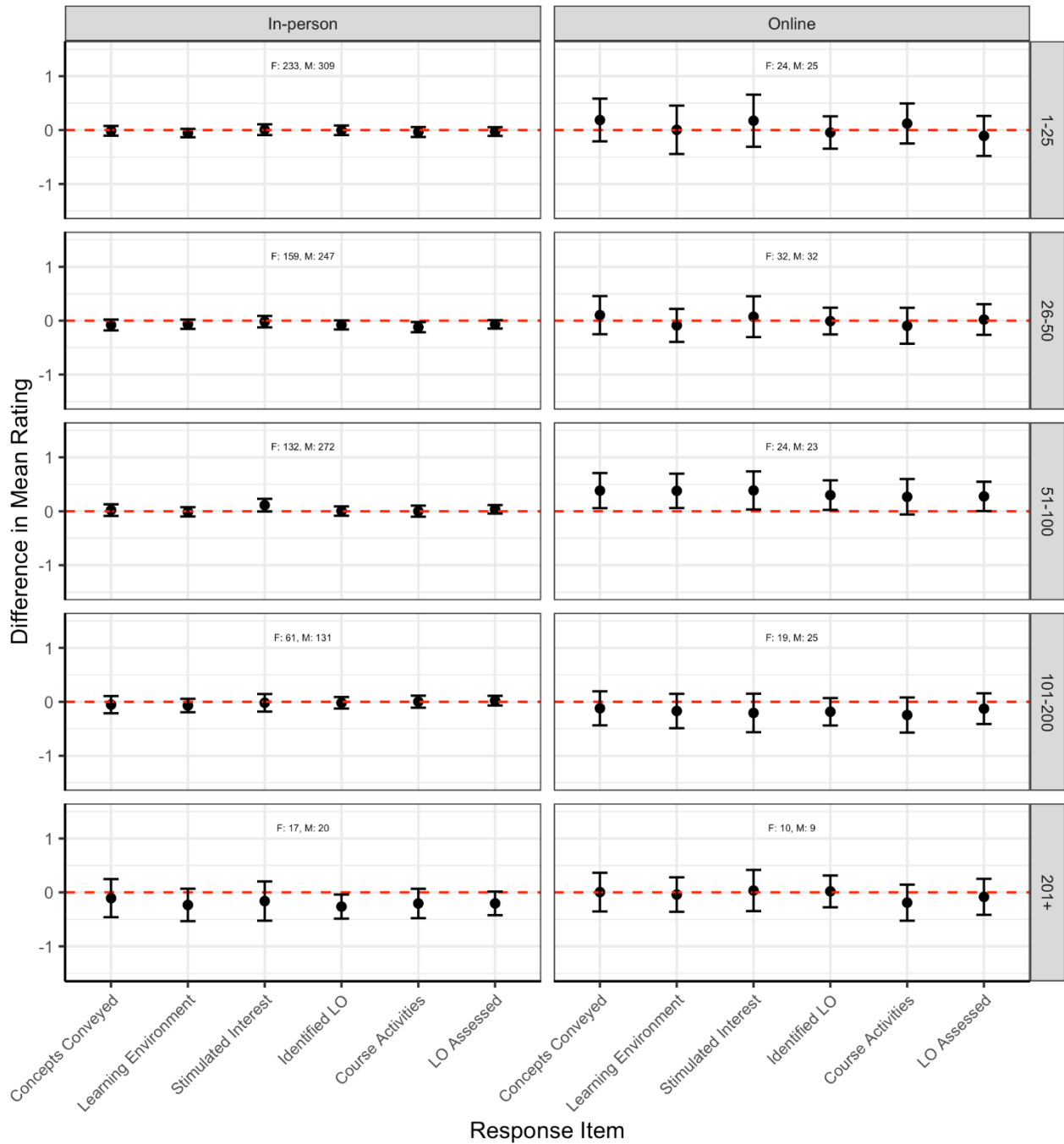


Figure 23: Difference in Winter 2025 mean rating for Male and Female Instructors by Course Type and Class Size. Error bars correspond to 95% confidence intervals

Recap of Key Findings

This report presents the University of Waterloo's fourth institutional analysis of Student Course Perceptions (SCP) survey data since the introduction of the new SCP instrument in Winter 2022. Drawing on 46,302 student responses across 2,483 section–instructor pairs, the Winter 2025 analysis reinforces a set of highly consistent patterns observed across prior years.

Several key findings emerge:

Stability of SCP results over time

- Across all six core items, SCP ratings remain strongly skewed toward the positive end of the five-point scale, with most responses falling within the “Agree” and “Strongly Agree” categories. This persistent pattern across four years of data provides strong evidence that the SCP is functioning as a stable and reliable instrument for capturing broad patterns in student perceptions of their learning experience.

Class size as the most consistent correlate of ratings

- Among all variables examined, class size continues to demonstrate the most consistent relationship with SCP scores. Smaller classes are associated with modestly higher ratings across all survey items, while larger classes tend to receive slightly lower ratings. Although these differences are consistent, they remain small in magnitude and should be interpreted cautiously. The pattern aligns with both prior Waterloo analyses and the broader literature on teaching and learning.

Minimal variation across Faculties and survey items

- Differences in mean ratings across Faculties remain tightly clustered within a narrow range, with only negligible variation on the five-point scale. Across contexts, items related to learning environment, clarity of concepts, and alignment of learning outcomes tend to be rated slightly higher than items related to stimulating interest and course activities, a pattern that has remained stable over time.

No evidence of systematic bias in SCP ratings between male and female instructors

- Consistent with findings since Winter 2022, analyses of differences in mean ratings assigned to male and female instructors (using HR sex-at-birth data) reveal no meaningful or systematic differences. Across the vast majority of comparisons, differences are not statistically significant; where statistically significant differences do appear, they are small, inconsistent in direction, and well below thresholds of practical significance. These results provide strong and sustained evidence that the SCP instrument, by design, does not produce gender-based disparities in student ratings at an institutional level.

Equity data limitations remain a critical constraint

- As in Winter 2024, high levels of nonresponse to the Equity Survey significantly limit the ability to conduct robust analyses related to racial identity, Indigenous identity, and gender identity. With more than half of instructors not providing equity data, subgroup analyses cannot be conducted without compromising reliability or confidentiality. As a result, equity-related findings remain descriptive only, representing a major gap in the institution's ability to examine potential sources of systemic bias.

Taken together, these findings underscore the consistency, stability, and interpretive boundaries of SCP data. Observed differences associated with instructor- and course-level variables are generally small and context-dependent, reinforcing the importance of situating SCP results within a broader, holistic framework for teaching assessment.

Concluding Remarks

The Winter 2025 analysis further strengthens confidence in the new Student Course Perceptions survey as a stable and thoughtfully designed instrument. Across four consecutive years of institutional analysis, SCP results have demonstrated remarkable consistency, not only in overall score distributions, but also in the relative relationships between items, class size, and instructor characteristics.

One of the most significant findings in the ongoing analyses of the SCP survey is the absence of systematic gender-based differences in SCP ratings. Since the introduction of the new instrument in Winter 2022, analyses have consistently found no evidence of statistically or practically meaningful disparities between male and female instructors. This sustained pattern is notable given longstanding concerns about gender bias in student evaluations of teaching and reflects the intentional design, validation, and ongoing monitoring of the SCP instrument.

At the same time, this analysis highlights a critical limitation that must be addressed to support the institution's equity goals: the lack of usable equity data. Without sufficient participation in the Equity Survey, it is not possible to rigorously examine whether SCP ratings vary across racialized, Indigenous, or other identity groups, or to identify potential sources of systemic bias. This represents a significant gap in the evidence base and constrains the institution's ability to fully evaluate the fairness and inclusivity of teaching assessment practices.

More broadly, the findings reinforce an important interpretive principle: small numerical differences in SCP scores, particularly on a compressed five-point scale, should not be overinterpreted. Even when statistically significant, differences of a few tenths of a point provide limited insight into teaching effectiveness in the absence of contextual information.

For this reason, SCP results are most meaningful when integrated with other sources of evidence. Open-ended student comments, peer review of teaching, and teaching dossiers provide essential context that helps explain patterns in quantitative ratings and supports more nuanced, fair, and development-oriented interpretations of teaching effectiveness.

Moving Forward: Implications for Teaching Assessment Practice

The Winter 2025 findings point to several important priorities for the continued evolution of teaching assessment at the University of Waterloo.

1. Strengthening participation in equity data collection

Increasing participation in the Equity Survey is an institutional priority. Without robust and representative equity data, it is not possible to assess the impact of systemic biases on SCP ratings. Addressing this challenge will require coordinated efforts to build trust, clarify the purpose and use of equity data, and ensure that data collection practices are perceived as appropriate, secure, and meaningful by instructors.

2. Sustaining confidence in the SCP instrument

The consistent absence of gender-based differences in SCP ratings across multiple years provides strong evidence of the instrument's consistency in this domain. Continued monitoring will remain important, particularly as additional data becomes available, but current findings support the use of SCP results as a credible component within the holistic teaching assessment model.

3. Interpreting SCP scores within context, not in isolation

The persistent clustering of scores at the high end of the scale and the small magnitude of observed differences reinforce the need for careful interpretation. SCP ratings are best understood as indicators of broad patterns rather than precise measures of teaching effectiveness. Greater emphasis should continue to be placed on integrating qualitative feedback and other forms of evidence to support meaningful interpretation and use.

4. Continuing to embed SCP within a holistic framework

Finally, findings reaffirm the importance of positioning SCP data within a multi-measure approach to teaching assessment. SCP ratings, when combined with student comments, peer review, and teaching dossiers, can support a more balanced and equitable understanding of teaching effectiveness, one that aligns with institutional commitments to educational quality, fairness, and continuous improvement.