

# Taking a Scholarly Approach to Transform Student Feedback on Teaching

---

Presented by:

**Stephanie McKeown, PhD, Abdel Azim  
Zumrawi, PhD, Alison Wong & Brad Wuetherick**

University of British Columbia

Sept 20<sup>th</sup>, 2023



**Course Evaluation Institute**



CENTRE FOR TEACHING SUPPORT & INNOVATION



# Acknowledge

---

The authors would like to acknowledge the following individuals who played a substantial role in the project assisting with data collection and analysis.

## **Academic Assistants (students)**

Anshul Dhariwal  
Lisvet Parra Montas  
Dyuti Raghu  
Thishani Rajapakshe

## **Offices of the Provosts and Vice Presidents Academic**

Debbie Hart (Vancouver)  
Laura Prada (Okanagan)

## **Office of Planning and Institutional Research**

Jodie Foster	Irene McKechnie
Brent Harris	Gavin Yap
Kristi Hoffman	

# The University of British Columbia

---



Course Evaluation Institute

# Outline

---

- **Background**
- **Methodology**
  - **Item Response Theory (IRT) Analysis**
  - **Differential Item Functioning (DIF) Analysis**
- **Sample Data**
- **Results**
- **Conclusions**
- **Next Steps**



# BACKGROUND

---

For additional details and updated information visit:

<https://seoi.ubc.ca/>



Course Evaluation Institute



# General Principles Recommended to Senates May 2020

---



Evaluations of Instruction at UBC will include Student Feedback through end-of-term course surveys



Name of the survey should be changed from 'Student Evaluation of Teaching' to 'Student experience of instruction' (SEI)



Questions asked of students should be student-focused and based on their experience with instructor(s) in specific contexts and relationships



Student leadership on both campuses should be actively engaged in raising the profile of SEI.

# Four Uses of Student Experience of Instruction Surveys

---

1. To provide data that will be used to continuously **improve the student's learning experience**.
2. To provide students, departments, faculties and the University with a source of data about the **overall quality of teaching**.
3. To provide teachers with information on their teaching performance and to assist with the **further development of their teaching**.
4. To provide the University with data on the quality of teaching to be used for operational purposes, including but not limited to **assessment of faculty for merit, promotion, tenure and institutional recognition**.

# Common University Questions for UBCV & UBCO

---

**Recommendation:** There should be a common set of university questions asked across both campuses – name of the survey should change from Student Evaluation of Teaching (SEOT) to Student Experience of Instruction (SEI) to be more student-centered.

**UBCV:** Changes to wording of questions 1, 2, 3, 5, 6; more significant changes to Q4

**UBCO:** Change from previous set of 19 questions to the same 6 core questions as UBCV



# Survey Items

---

- UBC experience of teaching survey items are modular in nature and include:
  - University Module Items (UMI);
  - Academic unit-specific items (Faculty/Department); and
  - Items that an instructor may add to the survey.
- The UMI and most academic unit-specific items use a 5-point Likert scale to indicate their level of agreement with the item statements:
  - 1= “Strongly Disagree”, 2= “Disagree”, 3= “Neutral”, 4= “Agree”, 5= “Strongly Agree”.

# Changes to UMI

---

## Old SET Questions

1. The instructor made it clear what students were expected to learn.
2. The instructor helped inspire interest in learning the subject matter.
3. The instructor communicated the subject matter effectively.
4. Overall, evaluation of student learning (through exams, essays, presentations, etc.) was fair.\*
5. The instructor showed concern for student learning.
6. Overall, the instructor was an effective teacher

## NEW SEI Questions

1. Throughout the term, the instructor explained course requirements so it was clear to me what I was expected to learn.
2. The instructor conducted this course in such a way that I was motivated to learn.
3. The instructor presented the course material in a way that I could understand.
4. Considering the type of class (e.g., large lecture, seminar, studio), the instructor provided useful feedback that helped me understand how my learning progressed during this course.
5. The instructor showed genuine interest in supporting my learning throughout this course.
6. Overall, I learned a great deal from this instructor.

# Student Experience of Instruction at UBC

---

- **Surveys are administered centrally by the PAIR Office for both campuses**
- **Supports both instructor and TA surveys**
- **Over 13,000 course sections, with more than 20,000 course-instructor/course-TA pairings in the Winter Session (September – April)**
- **General survey periods at the end of each term**
- **However, there are SEI surveys running all year long, e.g. off-cycle courses, sequential team-taught courses, rotations, practicums, etc.**

# Student Experience of Instruction at UBC

---

- **SEI surveys are to close before final exam**
- **Preparation and survey setup tasks starts about 4-6 weeks in advance:**
  - **We obtain data directly from upstream systems (Student Information System and HR Management System)**
  - **Working with faculty or department administrators to confirm courses and teaching assignments**
  - **Data checkpoints: faculty & department, instructor "heads-up" email, instructor survey notifications, student emails**
- **Instructor reports are available in**
  - **January (Winter Term 1),**
  - **May (Winter Term 2),**
  - **July (Summer Term 1) and August (Summer Term 2)**

# How we arrived at the new SEI questions

---





# METHODOLOGY

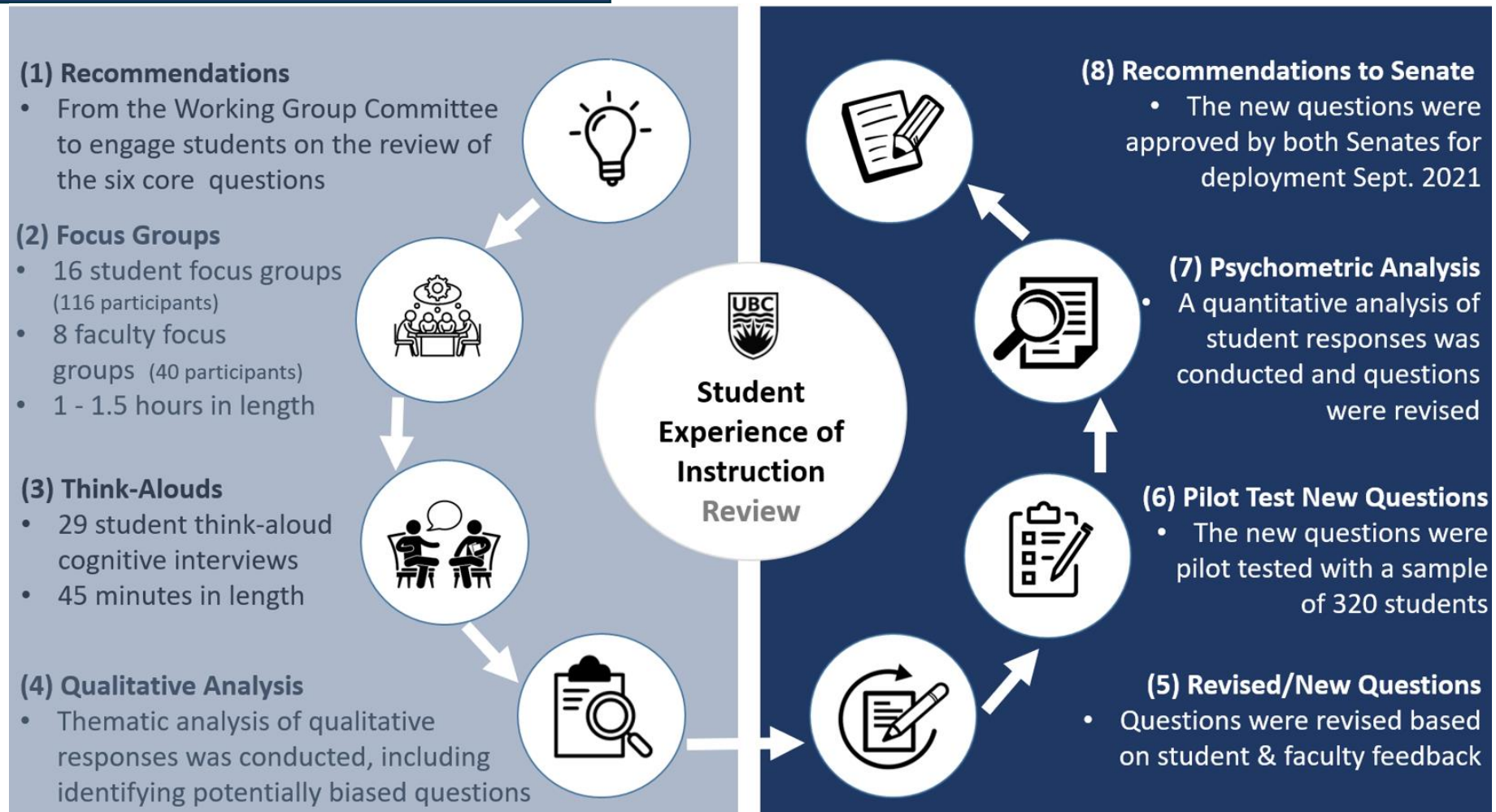
---



Course Evaluation Institute



# How we arrived at the new SEI questions



# Item Response Theory (IRT)

---

- IRT is an approach used for test development and can be used in a similar fashion for survey item development or refinement.
- Through IRT we are able to:
  - Predict individual survey responses based on a respondent's attitude or perception, and
  - Establish a relationship between an individual's item response and the set of traits underlying item performance by examining the Item Characteristic Curves.
- The information provided through an IRT analysis can help the survey developer evaluate how well the questions function across different attitudinal levels, and how well the response options work for each question.
- The analysis is focused on the interaction of the respondent with the individual items.



# Differential Item Functioning (DIF)

---

- DIF analyses can be used to examine whether students responded to the questions differently across groups, such as class size, campus, gender, etc. and can be used to identify bias.
- In surveys, DIF is conceptualized as occurring when survey respondents who have similar attitudes on a measured trait respond differently due to construct-irrelevant factors such as differential interpretation of terms used in the survey.
- Use multiple approaches and look for corresponding results: Mantel-Haenszel, Lord's Chi-square test (IRT-based) and regression-based methods (binary and ordinal).

# Differential Item Functioning (DIF)

---

- If an item is flagged as having DIF it suggests that a survey question may indicate a different understanding across the student groups.
- When DIF is detected, further analyses examine why some items function differentially across respondents to determine whether refinement of the survey question is needed.
- The magnitude of DIF can be determined as a) none or negligible, b) slight to moderate, and c) moderate to large (Zwick et al, 1999).



# SAMPLE DATA

---



Course Evaluation Institute



CENTRE FOR TEACHING SUPPORT & INNOVATION



# DATA – Testing The New Questions

---

- 2021 Winter term 1 survey data from both UBC campuses.
- 100 course surveys randomly selected from each of five fields of study

<u>Field of Study</u>	<u>Responses</u>
Engineering	1,892
Health Sciences	1,520
Humanities	1,784
Sciences	3,090
Social Sciences	2,746
Total	11,032

## Distribution of Sample Data

<u>Campus</u>	<u>Responses</u>
UBCO	2,134
UBCV	8,898

<u>Class Size</u>	<u>Responses</u>
< 100	4,519
>= 100	6,513
1 - 49	2,427
200+	2,891

\*Student gender captured from administrative records which were only recording binary gender fields at the time.

<u>Course Year Level</u>	<u>Responses</u>
1st Year	3,181
2nd Year	3,086
3rd Year	2,637
4th Year	969
5th Year	1,159
Total	11,032

<u>Student Gender*</u>	<u>Responses</u>
Female	6,542
Male	4,490

# Distribution of Sample Data

---

<u>Rank</u>	<u>Responses</u>
Assoc. Prof	1,845
Asst. Prof	2,917
Lecturer	1,754
Professor	1,933
Sessional	2,583

<u>Instructor Gender</u>	<u>Responses</u>
Female	4,211
Male	6,821

<u>Meeting time</u>	<u>Responses</u>
Early	3,635
Late	7,397



# RESULTS

---



Course Evaluation Institute





# IRT & DIF approaches

---

## Dimensionality Analysis



- ✓ Unidimensionality assumption  
Do all six UMIs measure a single underlying construct?
- ✓ Factor Analysis
- ✓ mirt package in R

## IRT Model Analysis



- ✓ Model Fit & Estimates
- ✓ Item Characteristics
- ✓ Test & Item Information
- ✓ mirt package in R

2-parameter multi-level IRT (MLRT) model

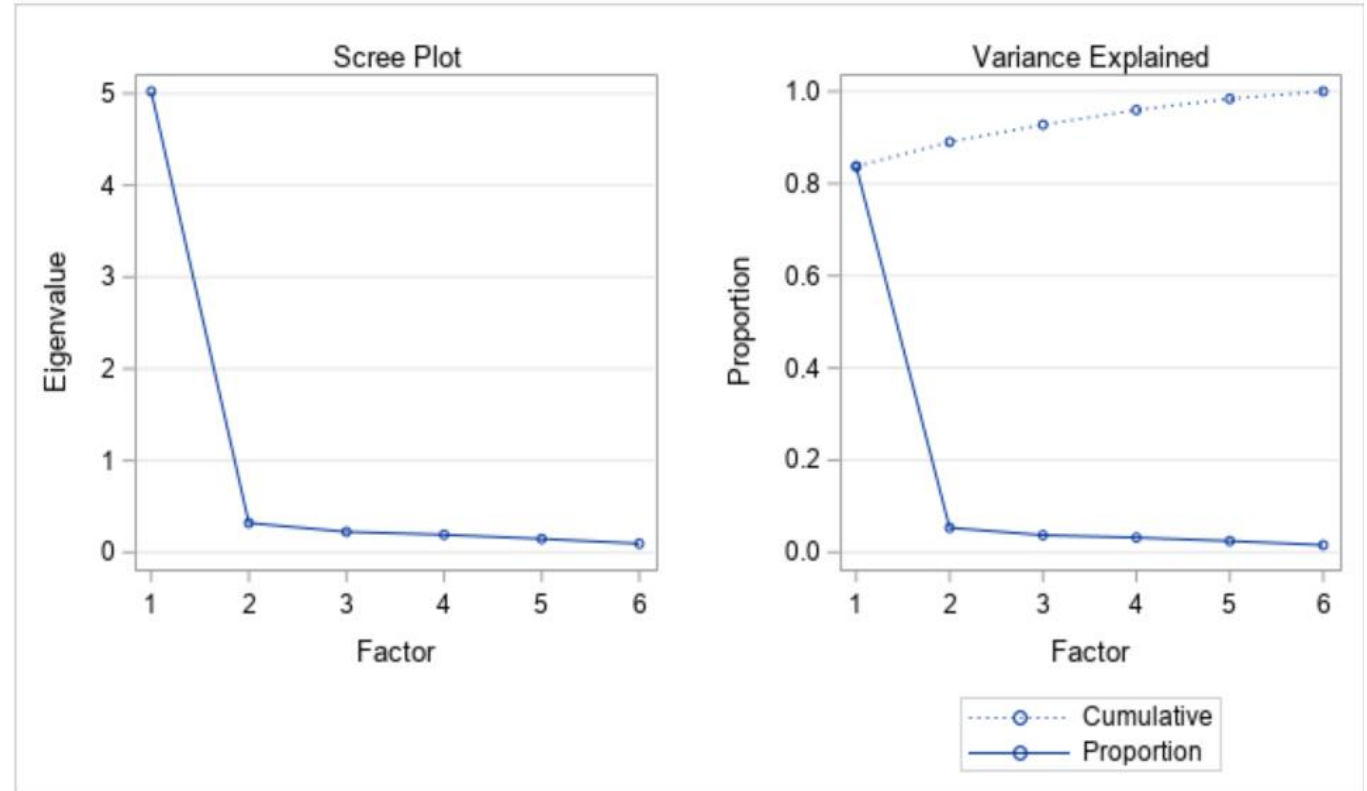
## Differential Item Functioning (DIF)



- ✓ Campus
- ✓ Year level
- ✓ Class size / Meeting time
- ✓ Instructor/Student gender
- ✓ mirt package in R
- ✓ LOGISTIC & GENMOD procedures in SAS

# Results - Dimensionality

- All six UMI items had high factor loadings, i.e. they represent a single underlying construct (student experience of instruction)
- Results of the factor analysis support the unidimensionality assumption for IRT analysis.

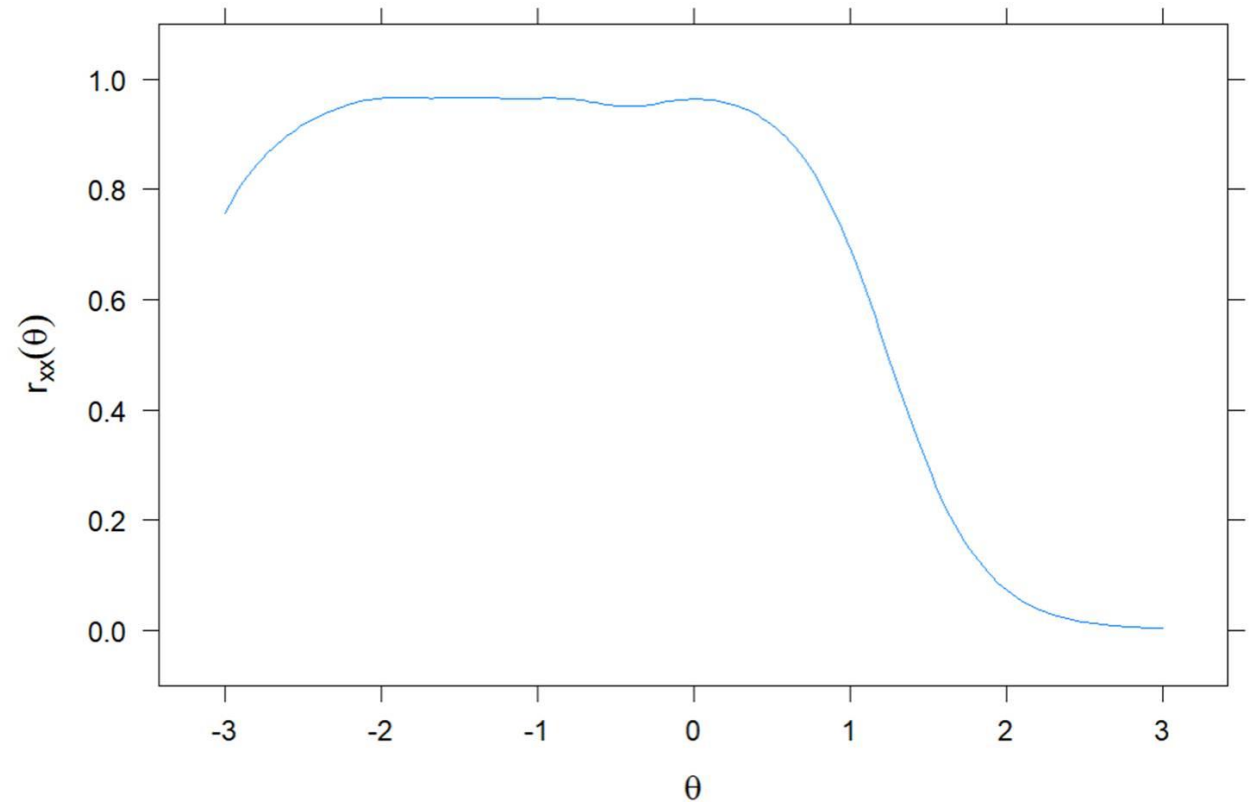


## Results - Comparing Models

Model	Criteria					$\chi^2$	df	p-value
	AIC	SABIC	HQ	BIC	logLik			
Base Model	112820.9	112944.8	112894.8	113040.2	-56380.46			
1-level	112617	112790.4	112720.4	112923.9	-56266.48	228	12	< 0.0001
MLRT	112883	113044.1	112979	113168	-56402.49			
1-level	112617	112790.4	112720.4	112923.9	-56266.48	272	3	< 0.0001

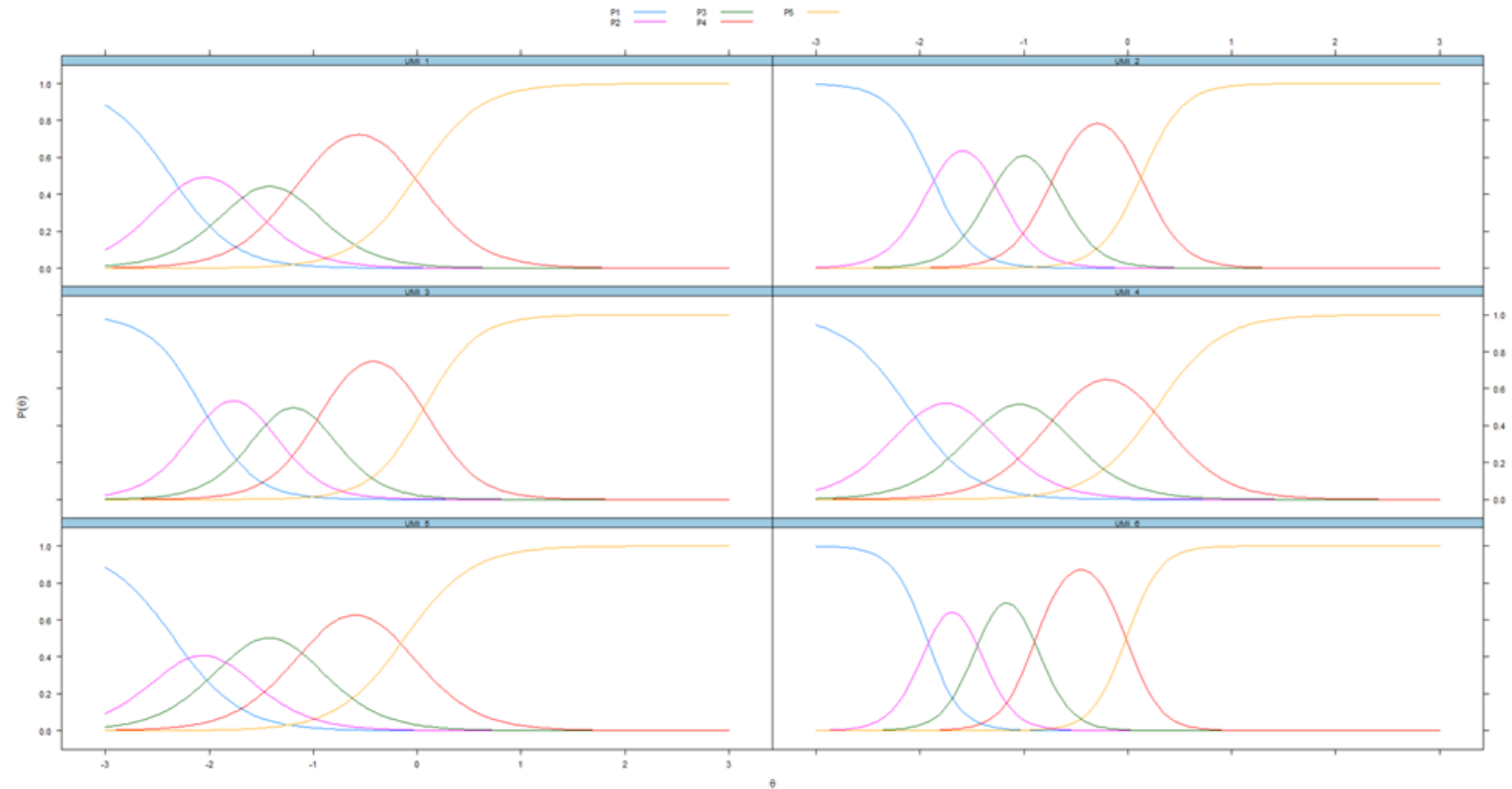
# Results - Reliability

- Cronbach's alpha: a measure of scale reliability, which indicates internal consistency. For the new/modified UMI items, Cronbach's alpha of 0.94 suggests a high survey reliability.
- Overall reliability of a survey based on:
  - how well the questions provide statistical information about the experience of instruction; and
  - how precisely scores can be estimated across different values of attitudinal scale
- An overall IRT marginal reliability estimate of 0.84, also suggests a high survey reliability.



# Results – Location Parameters

- Item location parameter estimates (thresholds) were fairly consistent across response options for the six UMI questions.
- The 5-point Likert scale options work well and function similarly within each of the six new UMI questions.





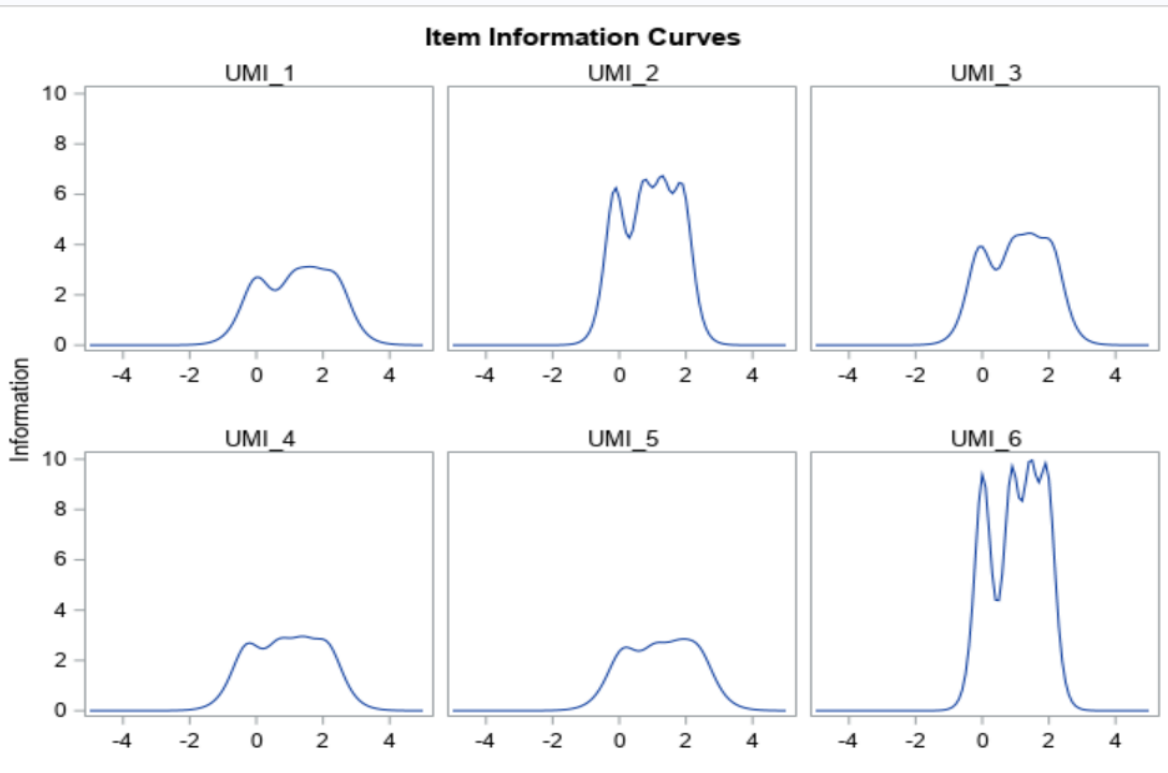
## Results – Discrimination Parameters

- Item discrimination parameter determines the rate at which the probability of positively endorsing an item change given the individual attitude/perception levels
- Typically, the larger the discrimination parameter, the steeper the slope, the more effective the item at discriminating among different attitudes along the continuum.

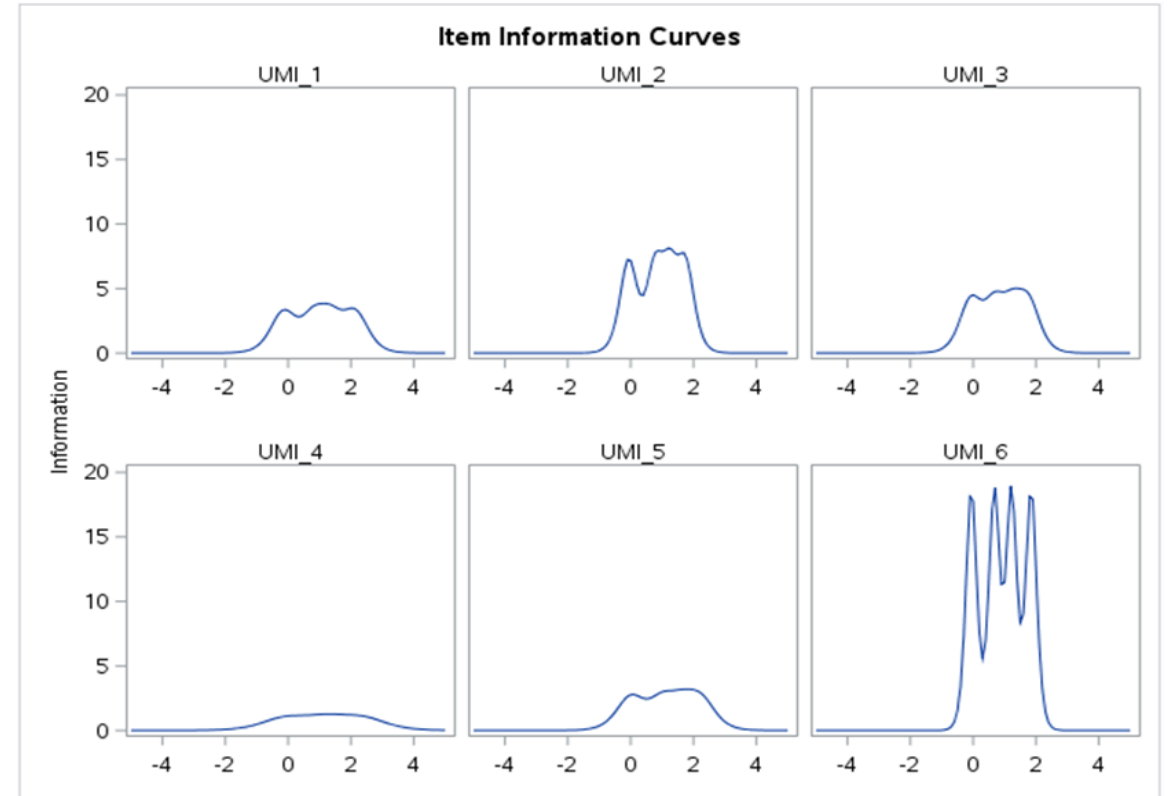
Data Source	Discrimination Parameter Estimates					
	UMI 1	UMI 2	UMI 3	UMI 4	UMI 5	UMI 6
Pre-2021 SEI Survey	3.62	5.38	4.15	2.02	3.28	8.67
2021 SEI Survey	3.26	4.80	3.83	3.15	3.00	5.85

# Results – Item Information Curves

## New UMI Questions



## Old UMI Questions



# Results – Differential Item Functioning (DIF)

	Grouping						
DIF Method	Campus	Student Gender	Class size < 100 vs > 100)	Class Size (1 – 49) vs 200+	Class Meeting Time	Year level 1 <sup>st</sup> , 2 <sup>nd</sup> & 3 <sup>rd</sup> vs 4 <sup>th</sup> & 5 <sup>th</sup>	Instructor Gender
Mantel-Haenszel*	Negligible	UMI 6 moderate	UMI 1 moderate	UMI 1, 4 (large) UMI 5, 6 moderate	Negligible	Negligible	UMI 3 moderate F
Logistic (Binary)**	None	UMI 6 uniform F	UMI 1 uniform >100	UMI 1, 4, 5, 6 uniform >50	None	None	UMI 3 uniform F
GLM (ordinal)**	---	UMI 6 uniform F	UMI 1 uniform >100	All uniform >50	None	----	UMI 3 uniform F
Lord's Chi-square Test	None	None	UMI 1	UMI 1, 2 & 6	None	None	UMI 3

# Conclusion

---

- Item Response Theory (IRT) results indicated that the new questions seem to function better than previous UMI questions.
- In the old version, UMI question 6 provided most of the statistical information for the overall survey.
- Item Information results provide further evidence that the new UMI questions are more consistent in their contribution to the overall survey
- Most of the new UMI questions showed no DIF among different grouping by student, instructor or class attributes

# Next Steps for the SEIs

---

- Quantitative Analysis
  - Obtain updated UBC Employment Equity Survey (EES) data (instructor gender; visible minority; indigenous; and disability status)
  - Combine the EES data with the 2021 & 2022 SEI survey data.
  - Use a hierarchical linear modelling approach (GLMM) to assess the effect of course attributes and instructor demographics on students' responses.
- Explore tools to support text analytics and visualization of qualitative feedback received through SEIs



# Faculty Use of SEIs

---

- The best instrument is only as valuable as our ability to support faculty (and academic leaders) to use the feedback:
  - Data literacy – Do faculty understand the data we report to them?
  - Feedback literacy – Can faculty use feedback effectively to improve their practices?
- Supporting faculty and academic leadership development, particularly through partnerships between PAIR and the Centre for Teaching and Learning (UBCO) and Centre for Teaching, Learning, and Technology (UBCV), is critical to faculty using student feedback effectively to achieve our stated purposes for SEIs

# SEIs as part of the Integrative Evaluation of Teaching

---

- As we continue to move forward with SEIs, it is critical to place them within the broader institutional priority around enhancing the Integrated Evaluation of Teaching, which at UBC includes, at a minimum, the following areas:
  - a. Peer Review of Teaching
  - b. Student Experience of Instruction
  - c. A statement of Self-Reflection
  
- At the fore of our governance conversations at UBC is an over-arching question – Against what standard are we evaluating teaching? (How are we defining effective / excellent teaching? What competencies are we expecting faculty / instructors to demonstrate through the sources of feedback received?)

# Questions / discussion

---



Course Evaluation Institute

