Effective Use of "Student Evaluation of Teaching" Data

PHIL GRANIERO AND LAURA WINER

INTERNATIONAL FORUM ON TEACHING EVALUATION

MAY 2, 2017

A rose by any other name…

Student Evaluation of Teaching (SET)

General Course Evaluation

Students' Evaluation of Education Quality (SEEQ)

Student End-of-Course Evaluation

SRI: Student Rating of Instruction (SRI)

SRI = SET = CE = SEEQ = …

Who are we? Why us?

That's Laura ←

That's Phil →

Getting to Know Each Other

What are your

- 'wish lists'
- questions
- challenges

with SET data?

bias from content
getting feedback later (the AHA moment)
response rates - how to increase
effective + responsible use of data
what data are collected?
linking to learning outcomes
different roles - instructor, reporting, system designers
use of aggregate vs. individual data

Aligning Questions, Data, and Goals

What is the Purpose of Course Evaluation?

Professor Chand's overall instructor score has averaged 5.6 over the last several years.

1	2	3	4	5	6	7
Extremely Poor	Very Poor	Poor	Adequate	Good	Very Good	Out-Standing

Most instructors hear: "Am I a good teacher?"

What is a better question?

What can Prof. Chand learn about her teaching based on the patterns found in her SET data?

A totally different way of thinking about the data.

- inquiry-based, improvement-based focus
- engages with the patterns, contradictions and challenges in the data
- open to reflection, follow-up questions, and "next steps"

This approach raises different questions

How has my teaching changed over time?

Are my strengths and weaknesses the same in every course?

What anomalies are there in my scores? Why?

Do my students respond better to me in some courses than in others?

What is the profile of what I've taught over time?

Are there other influences that should be documented?

Effective Use of SET Data

Use by whom?
- Instructor
- Tenure Committee
- Students
- Department Head
- Department Council

Use for what?
- Feedback / Communication
- Reflection / Practice
- Tenure/Promotion
- Workload Assignment
- Curriculum Development

Alignment with Institutional Goals

What questions people ask changes what they do.

Provision & reporting influences the kind of questions they ask.

People read a LOT into what you provide & report.

It drives what they think is important to you.

Regardless of whether it is what you actually think or not.

What are your SET goals?

What are you trying to get out of SETs?

What are you trying to use the scores for?

What are you trying to learn?

What questions are you trying to answer?

What decisions/actions are you trying to enable?

Understanding SET Scores

What *are* the numbers we're looking at?

University of Windsor

1	2	3	4	5	6	7
Extremely Poor	Very Poor	Poor	Adequate	Good	Very Good	Out-Standing

McGill

1	2	3	4	5
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

Do your response scales look something like this?

The numbers aren't really numbers...

1	2	3	4	5	6	7
Extremely Poor	Very Poor	Poor	Adequate	Good	Very Good	Out-Standing

...they just *look* like numbers.

The numbers are imprecise: fat chisel markers, not mechanical pencils.

Are average scores of 5.7 and 5.8 _really_ different?

Let's talk about averages...

An "average" implies that the class has one answer, but each student doesn't quite give the answer correctly.

An "average" implies that variations on each side effectively cancel each other out.

Is all that really true?

Same Means, Different Meaning

Same question, two different instructors

Consistently 'mediocre'
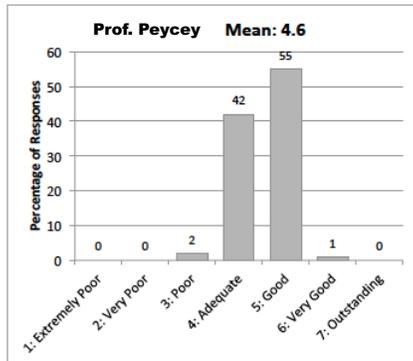
Polarized: either loved or hated

Prof. Peycey   Mean: 4.6

Percentage of Responses — values: 0, 0, 2, 42, 55, 1, 0 across categories 1: Extremely Poor, 2: Very Poor, 3: Poor, 4: Adequate, 5: Good, 6: Very Good, 7: Outstanding

Prof. Mac   Mean: 4.6

Percentage of Responses — values: 10, 17, 15, 5, 3, 17, 36 across categories 1: Extremely Poor, 2: Very Poor, 3: Poor, 4: Adequate, 5: Good, 6: Very Good, 7: Outstanding

---
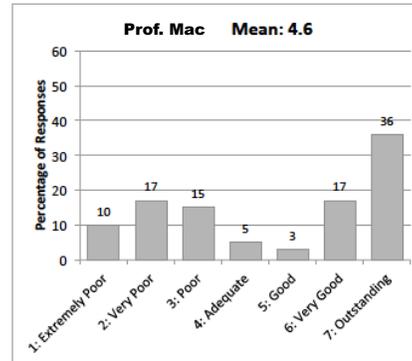
Our students aren't identical...

...so why over-simplify their different experiences with a single number?

Useful Noise....

SET captures the variety of experience in your classroom.

The **distribution** is <u>*very*</u> informative.

Averages try to smooth out 'background noise'.

Averages try to condense and summarize, which *is* helpful.

But with SET scores, they *also* remove a lot of the message.

Make sure you can also look at the noise.
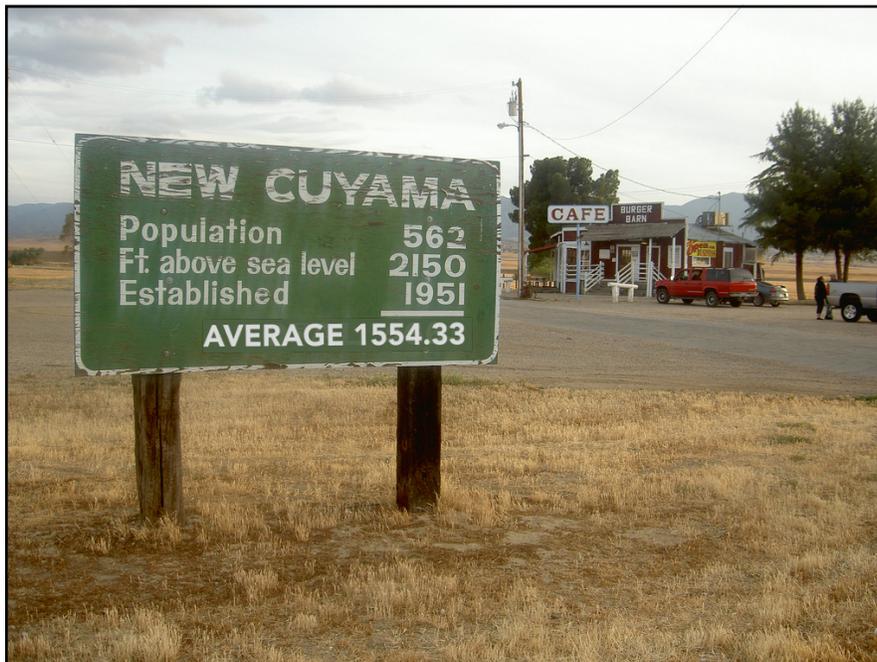
Speaking of averages...

What does an overall, average score represent about our multi-faceted teaching?

Don't ignore the individual questions!

Report Formatting Matters

REPORT #1 University of Windsor January 12, 2013

Please distribute to appropriate instructors

Student Evaluation of Instructors by Course/section taught Fall 2012

Course: 9900101   Section: 01 Enrollment:41   Courses:1   (all courses)

Instructor: CARVEY, D. Forms Scanned:28(=68% of enrollment)

A. Questions about the Instructor

Question | # of Responses to Scale Rating

	1	2	3	4	5	6	7	NA	No.	Avg.
A1	0	3	1	3	13	7	1	0	28	4.8
A2	2	1	3	2	8	7	4	0	27	4.9
A3	1	0	2	1	8	10	6	0	28	5.5
A4	1	0	0	1	6	13	7	0	28	5.8
A5	0	0	0	0	6	9	12	0	27	6.2
A6	0	0	2	2	8	11	5	0	28	5.5
A7	0	0	0	2	6	10	10	0	28	6.0
A8	0	0	0	0	5	9	13	1	27	6.3
A9	0	0	0	0	8	10	6	4	24	5.9
A10	0	0	0	0	8	13	6	0	27	5.9
A11	0	0	0	0	8	9	9	0	26	6.0
A12	0	1	3	0	7	10	5	0	26	5.4

Weighted Average Instructor Score 5.7

Here's what your report might look like.

It doesn't have to be pretty to be better...

Question	# of Responses to Scale Rating							NA	No.	Avg.
	1	2	3	4	5	6	7			
A1		3	1	3	13	7	1		28	4.8
A2	2	1	3	2	8	7	4		27	4.9
A3	1		2	1	8	10	6		28	5.5
A4	1			1	6	13	7		28	5.8
A5					6	9	12		27	6.2
A6		2	2		8	11	5		28	5.5
A7			2		6	10	10		28	6.0
A8					5	9	13	1	27	6.3
A9					8	10	6	4	24	5.9
A10					8	13	6		27	5.9
A11					8	9	9		26	6.0
A12		1	3		7	10	5		26	5.4

Weighted Average Instructor Score 5.7

...but let's make it prettier anyways.

Instructor related questions:	Mean	Extremely Poor	Very Poor	Poor	Adequate	Good	Very Good	Out-Standing
1. presented material in an organized, well-planned manner	4.8		3	1	3	13	7	1
2. was approachable for additional help	5.9					8	13	6
3. was accessible to students for individual consultation (in office hours, after class, open-door, by e-mail,phone)	6					8	9	9
4. The overall effectiveness of the instructor was	5.4		1	3		7	10	5
5. used instructional time well	4.9	2	1	3	2	8	7	4
6. explained content clearly with appropriate use of examples	5.5	1		2	1	8	10	6
7. was a clear and effective speaker	5.8	1			1	6	13	7
8. communicated enthusiasm and interest in the course material	6.2					6	9	12
9. stimulated your interest in the subject and motivated your learning	5.5			2	2	8	11	5
10. attended to students' questions and answered them clearly and effectively	6				2	6	10	10
11. was open to students' comments and suggestions	6.3					5	9	13
12. was sensitive to students' difficulties	5.9					8	10	6

What was the variation of experience within the classroom?

Instructor related questions:	Graph	Mean	Extremely Poor	Very Poor	Poor	Adequate	Good	Very Good	Out-Standing
1. presented material in an organized, well-planned manner		4.8		3	1	3	13	7	1
2. was approachable for additional help		5.9					8	13	6
3. was accessible to students for individual consultation (in office hours, after class, open-door, by e-mail,phone)		6					8	9	9
4. The overall effectiveness of the instructor was		5.4		1	3		7	10	5
5. used instructional time well		4.9	2	1	3	2	8	7	4
6. explained content clearly with appropriate use of examples		5.5	1		2	1	8	10	6
7. was a clear and effective speaker		5.8	1			1	6	13	7
8. communicated enthusiasm and interest in the course material		6.2					6	9	12
9. stimulated your interest in the subject and motivated your learning		5.5			2	2	8	11	5
10. attended to students' questions and answered them clearly and effectively		6				2	6	10	10
11. was open to students' comments and suggestions		6.3					5	9	13
12. was sensitive to students' difficulties		5.9					8	10	6

REPORT #1   **University of Windsor**   January 12, 2013

Please distribute to appropriate instructors

Student Evaluation of Instructors by Course/section taught   Fall 2012

Course: 9900101 Section: 01   Enrollment:41   Courses:1   (all courses)

Instructor: CARVEY, D.   Forms Scanned:28(=68% of enrollment)

A. Questions about the Instructor

Question   # of Responses to Scale Rating

	1	2	3	4	5	6	7	NA	No.	Avg.
A1	0	3	1	3	13	7	1	0	28	4.8
A2	2	1	3	2	8	7	4	0	27	4.9
A3	1	0	2	1	8	10	6	0	28	5.5
A4	1	0	0	1	6	13	7	0	28	5.8
A5	0	0	0	0	6	9	12	0	27	6.2
A6	0	0	2	2	8	11	5	0	28	5.5
A7	0	0	0	2	6	10	10	0	28	6.0
A8	0	0	0	0	5	9	13	1	27	6.3
A9	0	0	0	0	8	10	6	4	24	5.9
A10	0	0	0	0	8	13	6	0	27	5.9
A11	0	0	0	0	8	9	9	0	26	6.0
A12	0	1	3	0	7	10	5	0	26	5.4

Weighted Average Instructor Score 5.7

You can see all that here, right?

It doesn't have to be pretty to be better...

Question	# of Responses to Scale Rating							NA	No.	Avg.
	1	2	3	4	5	6	7			
A1		3	1	3	13	7	1		28	4.8
A2	2	1	3	2	8	7	4		27	4.9
A3	1		2	1	8	10	6		28	5.5
A4	1			1	6	13	7		28	5.8
A5					6	9	12		27	6.2
A6		2	2		8	11	5		28	5.5
A7			2		6	10	10		28	6.0
A8					5	9	13	1	27	6.3
A9					8	10	6	4	24	5.9
A10					8	13	6		27	5.9
A11					8	9	9		26	6.0
A12		1	3		7	10	5		26	5.4

Weighted Average Instructor Score 5.7

...and it doesn't have to be complicated. This is an Excel template.

Instructor related questions:	Graph	Mean	Extremely Poor	Very Poor	Poor	Adequate	Good	Very Good	Out-Standing
1. presented material in an organized, well-planned manner		4.8		3	1	3	13	7	1
2. was approachable for additional help		5.9					8	13	6
3. was accessible to students for individual consultation (in office hours, after class, open-door, by e-mail,phone)		6					8	9	9
4. The overall effectiveness of the instructor was		5.4		1	3		7	10	5
5. used instructional time well		4.9	2	1	3	2	8	7	4
6. explained content clearly with appropriate use of examples		5.5	1		2	1	8	10	6
7. was a clear and effective speaker		5.8	1			1	6	13	7
8. communicated enthusiasm and interest in the course material		6.2					6	9	12
9. stimulated your interest in the subject and motivated your learning		5.5			2	2	8	11	5
10. attended to students' questions and answered them clearly and effectively		6				2	6	10	10
11. was open to students' comments and suggestions		6.3					5	9	13
12. was sensitive to students' difficulties		5.9					8	10	6

Core Reporting Principles

The format of the report has a strong influence on:
- what questions occur to people.
- what data are interpretable and usable.
- what meaningful insights are noticeable.
- what inspires people to inquire further.

Make the data you are reporting as flexible and extendable as possible

Take out as much of the immaterial data/noise as possible, while leaving in or highlighting as much of the meaningful data/variation as possible.

Eliminate as many numerically/statistically invalid or misleading measures as you can. If you must keep them, complement them with other, more appropriate measures and (preferably visual) aids to guide readers towards meaningful interpretations.

Zoom
You are doing it wrong...

Here's another question:

Professor Carvey's overall instructor score has averaged 5.6 over the last several years.

1	2	3	4	5	6	7
Extremely Poor	Very Poor	Poor	Adequate	Good	Very Good	Out-Standing

Is he a good teacher?

From _one number_? Really?

Averages imply that there might be some variation, but things basically stay the same.

Dr. Carvey - Mean Instructor Scores 2007/08 to 2014/15

Same Means, Different Meaning

Steady Growth

Crash

Steady Decline
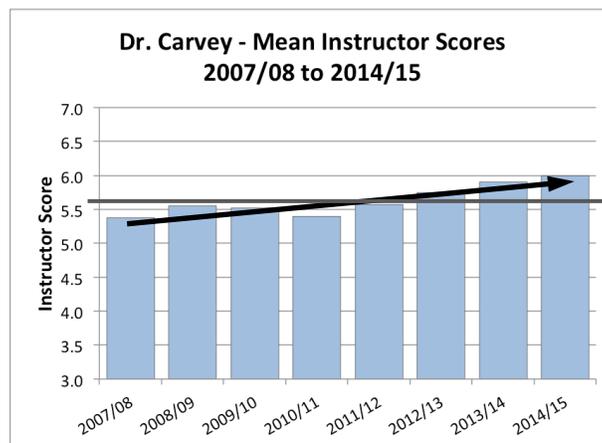
Mixed Bag
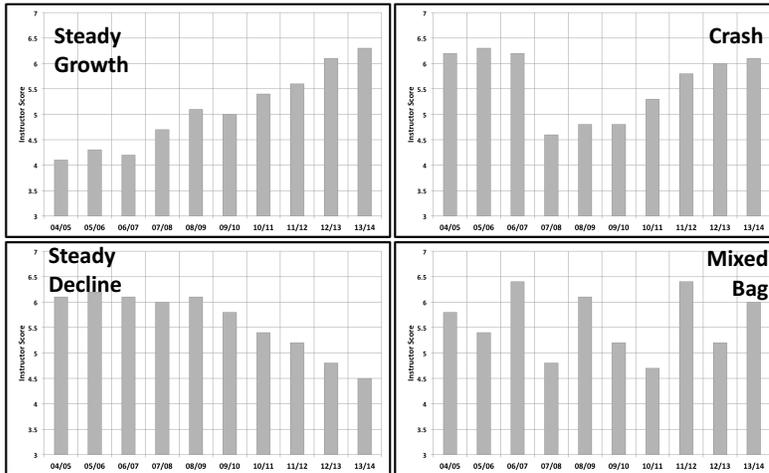
These all have an overall mean of 5.6.

Is the story the same for each one?

And what about all the courses?

And what about all the questions?

Live Demo!

SET Data Aren't Just Numbers

An Inquiry-Based Approach

Prof. Ferrell feels strongly that how the students responded to some questions completely contradicts what she believes she is actually doing.

◦ *"I do double the number of office hours compared to anyone in my department, I stay after class, but they keep saying I'm not accessible. It was never a problem in the past, but for the past couple of years I'm getting hammered."*

◦ *"I have all kinds of charts for them that outline the course and timeline, I give them planning guides for the midterm and the final, and I give them the PowerPoint slides and other prep material well in advance. It's find in some of my courses, but in the others I get slammed on organization every year."*

Where is the disconnect?

Written comments help a lot!!!

Comments Analysis Worksheet					
Comment Category	Sample Positive Student Comments	Total +	Sample Negative Student Comments	Total -	Personal notes
Overall (Course or Instructor)					
Clarity & Difficulty					
Organization & Structure					
Interest					
Teaching Strategies					
Assessment & Feedback					
Outside of Classroom Communications					
Personal Traits					
Physical Environment					

SET for Departments

Live Demo!

Ideas Check-In

What new questions do you have?

What new possibilities do you see?

Condensed Wisdom

Key Lesson Learned

> **The main problems that need solving are _human_ in nature, not technical.**

> **The inevitable, on-going challenge of engagement in autonomous, decentralized academic cultures.**

Wisdom Theme #1

It turns out that our primary task isn't figuring out how to structure the data.

**Our task is first figuring out what kind of *thinking* people need to do,
then figure out how to represent the data so they *can* think that way.**

Wisdom Theme #2

The sooner the data don't look like numbers, the better.

**When our prototypes focused on numerical representations, we got nowhere.
When our prototypes focused on reports, we made little progress.**

**When our prototypes focused on graphical representations, we made huge leaps forward.
When our prototypes focused on data interaction, we made huge leaps forward.**

Wisdom Theme #3

Institutional change is hard: interests are always at odds among different levels and different stakeholders.

**Imposed change generally fails to gain traction.
People need to see they have something to gain before buying in.**

Grassroots change is engaged, but difficult to scale up.

Usually, the tools that evolve successfully are the ones that every stakeholder group at every level finds ways to use them to further their interests.

Putting it to work

What *simple* thing might you try *first*?
- using your current reports
- modifying your current reports

And what after *that*?
- next steps
- concerns to address

Thank you kindly.

If you have questions later, are seeking friendly advice, or wish to have a chat:

Laura Winer: laura.winer@mcgill.ca

Phil Graniero: graniero@uwindsor.ca

International Forum on Teaching Evaluation

Windsor-Oakland International Teaching & Learning Conference

http://ctl2.uwindsor.ca/tclconf/