

# Investigating the influence of instructions and items on student evaluations of teaching

Kate A. Ratliff  
University of Florida  
Department of Psychology

Collaborators (in alphabetical order):

John Jordi, Nicole Lofaro, David Miller, Bonnie Moradi, Amanda Phalin, Christine Vitiello

# Student Evaluations of Teaching (SETs)

---

Surveys administered to enrolled students by colleges and universities at the end of an academic term

- Have been used since the early 20<sup>th</sup> century to assess students' impressions
- Are frequently interpreted as a metric of instructor quality
- Have come under increasing scrutiny and widespread criticism...

SETs are overly sensitive to information irrelevant to student learning

- Expected grade (Benton & Li, 2015)
- Class level (Davies et al., 2007)
- Instructor demographics (e.g., race, gender; Boring, 2017; Mengal et al., 2018)

# Background

---

Best practices for SETs have been proposed—but not widely tested

- *SET instructions*: What students read prior to completing SETs
- *SET items*: The questions to which students respond when completing SETs)

## **Bluenotes Faculty Grant Specific Aim**

Investigate whether implementing best practices for SET instructions and SET questions impacts faculty and course ratings on student evaluations of teaching

# Overview of Proposed Studies

---

## Study 1

### **Naturalistic, Field Study SET Instructions**

Compare average SETs from Spring 2020 to those from Spring 2019, before best practice instructions were implemented university-wide at the University of Florida

## Study 2

### **Experimental, Lab Study SET Items**

Participants evaluate a fictitious course and complete SETs that use either standard SET items or best practice SET items

## Study 3 (Future)

### **Experimental, Field Study SET Items**

Using what we learn in Study 2, UF students will be randomly assigned to complete actual SETs that use either standard SET item or best practice SET items



# Overview of Actual Studies

## Study 1

**Naturalistic, Field Study  
SET Instructions**

Compare average SETs from Spring 2020 to those from Spring 2019, before best practice instructions **AND FULLY ONLINE COURSES** were implemented university-wide at the University of Florida

## Study 2

**Experimental, Lab Study  
SET Items**

Participants evaluate a fictitious course and complete SETs that use either standard SET items or best practice SET items  
**MOVED STUDY ONLINE, ABBREVIATED**

## Study 3 (Future)

**Experimental, Field Study  
SET Items**

Using what we learn in Study 2, UF students will be randomly assigned to complete actual SETs that use either standard SET item or best practice SET items  
**ON HOLD**

# Study 1 Overview, Course Sample, and Participants

## Study 1 Naturalistic, Field Study SET Instructions

Compare average SETs from Spring 2020 to those from Spring 2019, before best practice instructions **AND FULLY ONLINE COURSES** were implemented university-wide at the University of Florida

**N = 391** courses with the same course name, number, instructor, and format in Spring 2019 and Spring 2020

--

7 Colleges

16 Departments

--

282 unique faculty members

--

267 unique courses

--

Replicated with data comparing Fall 2019 to Spring 2020, N = 308 courses

### Instructor Gender

Men	44%
Women	56%

### Instructor Race

White	74%
Non-White	26%

### Instructor Tenure Status

Tenured	23%
Non-Tenured	77%

### Course Level

Undergrad	63%
Grad	37%

### Course Format

Primarily In-Person	55%
Primarily Online	38%

# University of Florida SET Items

---

Item	<i>Mean (SD)</i>
------	------------------



# University of Florida SET Items

---

Item	<i>Mean (SD)</i>
The instructor was enthusiastic about the course.	4.61 (0.35)

# University of Florida SET Items

---

Item	<i>Mean (SD)</i>
The instructor was enthusiastic about the course.	4.61 (0.35)
The instructor explained material clearly and in a way that enhanced my understanding	4.37 (0.46)

# University of Florida SET Items

Item	Mean (SD)
The instructor was enthusiastic about the course.	4.61 (0.35)
The instructor explained material clearly and in a way that enhanced my understanding	4.37 (0.46)
The instructor maintained clear standards for response and availability	4.50 (0.39)

# University of Florida SET Items

Item	Mean (SD)
The instructor was enthusiastic about the course.	4.61 (0.35)
The instructor explained material clearly and in a way that enhanced my understanding	4.37 (0.46)
The instructor maintained clear standards for response and availability	4.50 (0.39)
The instructor fostered a positive learning environment that engaged students	4.50 (0.39)

# University of Florida SET Items

Item	Mean (SD)
The instructor was enthusiastic about the course.	4.61 (0.35)
The instructor explained material clearly and in a way that enhanced my understanding	4.37 (0.46)
The instructor maintained clear standards for response and availability	4.50 (0.39)
The instructor fostered a positive learning environment that engaged students	4.50 (0.39)
The instructor provided prompt and meaningful feedback	4.38 (0.48)

# University of Florida SET Items

Item	Mean (SD)
The instructor was enthusiastic about the course.	4.61 (0.35)
The instructor explained material clearly and in a way that enhanced my understanding	4.37 (0.46)
The instructor maintained clear standards for response and availability	4.50 (0.39)
The instructor fostered a positive learning environment that engaged students	4.50 (0.39)
The instructor provided prompt and meaningful feedback	4.38 (0.48)
The instructor was instrumental to my learning in the course	4.34 (0.50)

# University of Florida SET Items

Item	Mean (SD)
The instructor was enthusiastic about the course.	4.61 (0.35)
The instructor explained material clearly and in a way that enhanced my understanding	4.37 (0.46)
The instructor maintained clear standards for response and availability	4.50 (0.39)
The instructor fostered a positive learning environment that engaged students	4.50 (0.39)
The instructor provided prompt and meaningful feedback	4.38 (0.48)
The instructor was instrumental to my learning in the course	4.34 (0.50)
The course was relevant and useful	4.37 (0.38)

# University of Florida SET Items

Item	Mean (SD)
The instructor was enthusiastic about the course.	4.61 (0.35)
The instructor explained material clearly and in a way that enhanced my understanding	4.37 (0.46)
The instructor maintained clear standards for response and availability	4.50 (0.39)
The instructor fostered a positive learning environment that engaged students	4.50 (0.39)
The instructor provided prompt and meaningful feedback	4.38 (0.48)
The instructor was instrumental to my learning in the course	4.34 (0.50)
The course was relevant and useful	4.37 (0.38)
The course fostered regular interaction between the instructor and the students	4.23 (0.52)



# University of Florida SET Items

Item	Mean (SD)
The instructor was enthusiastic about the course.	4.61 (0.35)
The instructor explained material clearly and in a way that enhanced my understanding	4.37 (0.46)
The instructor maintained clear standards for response and availability	4.50 (0.39)
The instructor fostered a positive learning environment that engaged students	4.50 (0.39)
The instructor provided prompt and meaningful feedback	4.38 (0.48)
The instructor was instrumental to my learning in the course	4.34 (0.50)
The course was relevant and useful	4.37 (0.38)
The course fostered regular interaction between the instructor and the students	4.23 (0.52)
Course activities and assignments improved my ability to analyze, solve problems, and/or think critically	4.31 (0.41)

# University of Florida SET Items

Item	Mean (SD)
The instructor was enthusiastic about the course.	4.61 (0.35)
The instructor explained material clearly and in a way that enhanced my understanding	4.37 (0.46)
The instructor maintained clear standards for response and availability	4.50 (0.39)
The instructor fostered a positive learning environment that engaged students	4.50 (0.39)
The instructor provided prompt and meaningful feedback	4.38 (0.48)
The instructor was instrumental to my learning in the course	4.34 (0.50)
The course was relevant and useful	4.37 (0.38)
The course fostered regular interaction between the instructor and the students	4.23 (0.52)
Course activities and assignments improved my ability to analyze, solve problems, and/or think critically	4.31 (0.41)
Overall, this course was a valuable educational experience.	4.38 (0.41)

# University of Florida SET Items

Item	Mean (SD)
The instructor was enthusiastic about the course.	4.61 (0.35)
The instructor explained material clearly and in a way that enhanced my understanding	4.37 (0.46)
The instructor maintained clear standards for response and availability	4.50 (0.39)
The instructor fostered a positive learning environment that engaged students	4.50 (0.39)
The instructor provided prompt and meaningful feedback	4.38 (0.48)
The instructor was instrumental to my learning in the course	4.34 (0.50)
The course was relevant and useful	4.37 (0.38)
The course fostered regular interaction between the instructor and the students	4.23 (0.52)
Course activities and assignments improved my ability to analyze, solve problems, and/or think critically	4.31 (0.41)
Overall, this course was a valuable educational experience.	4.38 (0.41)
<b>Overall SET Score (N = 391); <math>\alpha = .96</math></b>	<b>4.40 (0.37)</b>

# Predicting Average SET Scores

Spring 2019

Simultaneous Regression					Simple Effects		
Predictor	B	$\beta$	SE	<i>p</i>	Group 0 <i>M</i> (SD)	Group 1 <i>M</i> (SD)	<i>d</i>
<b>Instructor Gender</b> 0 = Man, 1 = Woman	.028	.036	.040	.482	4.38 (0.36)	4.41 (0.38)	0.08
<b>Instructor Race</b> 0 = White, 1 = Non-White	-.058	-.066	.046	.204	4.42 (0.38)	4.34 (0.34)	0.22
<b>Instructor Tenure Status</b> 0 = Tenure, 1 = NT	-.034	-.038	.046	.464	4.45 (0.33)	4.39 (0.38)	0.17
<b>Course Level</b> 0 = UG, 1 = G	.194	.243	.042	< .0001	4.47 (0.38)	4.30 (0.38)	0.45
<b>Course Format</b> 0 = In-person, 1 = Online	-.207	-.264	.042	< .0001	4.34 (0.39)	4.50 (0.32)	0.45

# PROVIDING CONSTRUCTIVE FEEDBACK

Teaching is fundamental to the University of Florida. We appreciate your confidential feedback, which improves learning and is also used for Faculty development and evaluation.

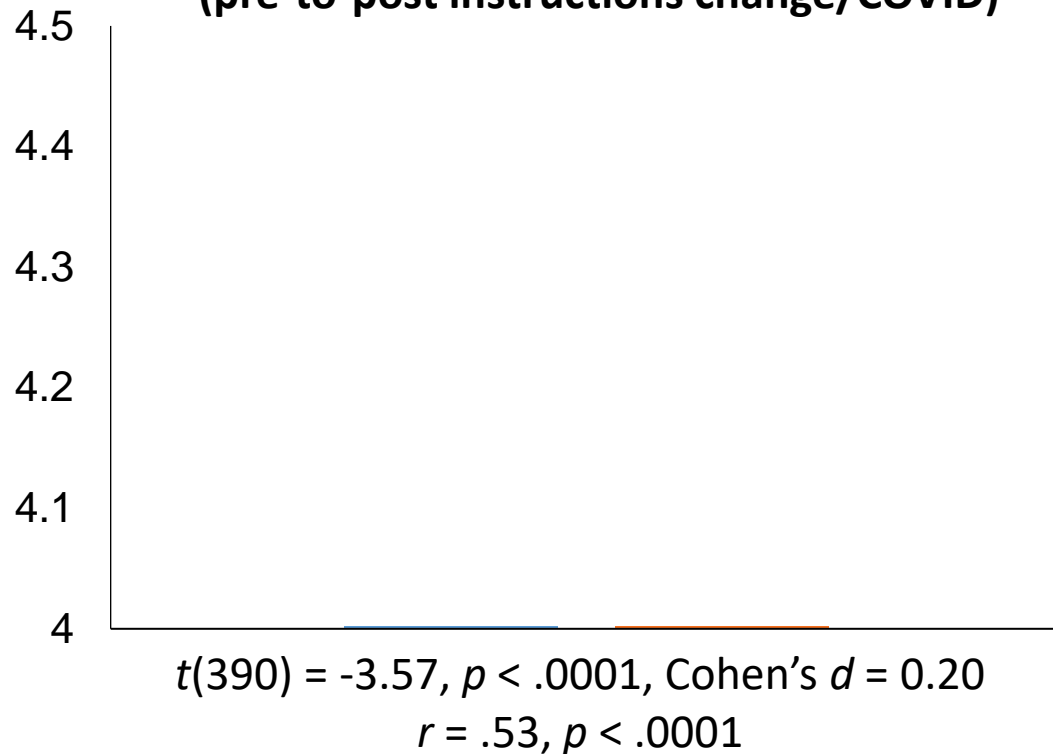
Research shows that we may evaluate identical course content and teaching differently based on unintentional stereotypes, such as professor age, gender, nationality, race, title or other characteristics. For example, women and people of color may be rated lower even when no actual differences exist.

Please keep this in mind and focus evaluations on course content (assignments, readings, in-class materials) and not unrelated matters (instructor's appearance).

<https://gatorevals.aa.ufl.edu/students/>

# Pre-Post Comparison: Average SET Scores

## Spring 2019 to Spring 2020 Comparison (pre-to-post instructions change/COVID)



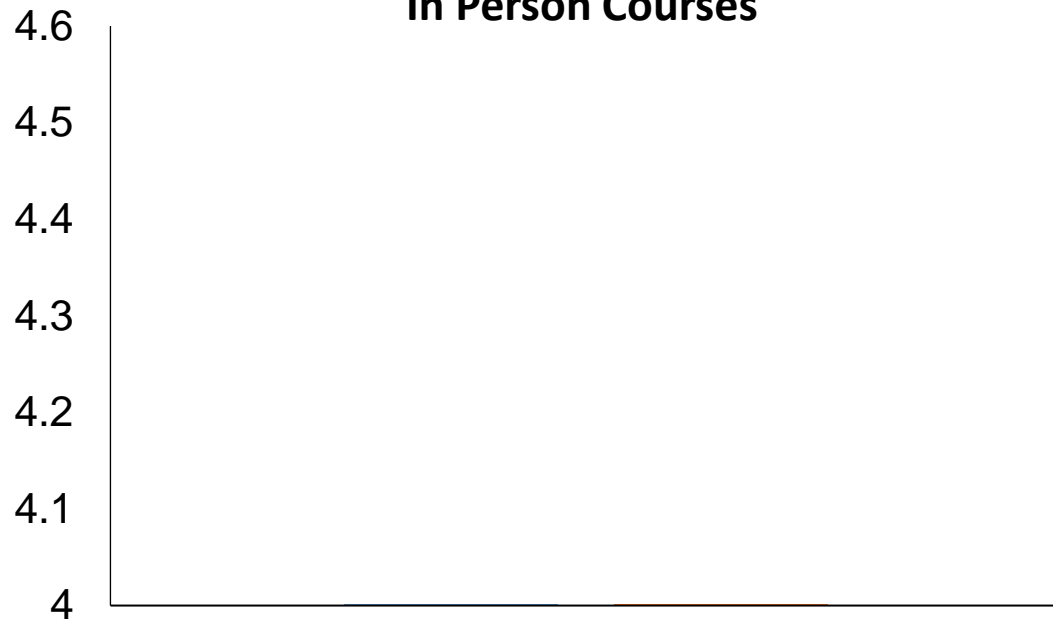
### The instructor...

- was enthusiastic about the course\*
- explained material clearly in a way that enhanced understanding\*
- maintained clear standards for response and availability\*
- fostered a positive learning environment that engaged students\*
- provided prompt and meaningful feedback\*
- was instrumental to my learning in the course\*

- Course content was relevant and useful
- Course fostered regular interaction
- Course activities and assignments improved my ability to analyze, solve problems, and/or think critically.
- Overall, this course was a valuable educational experience.

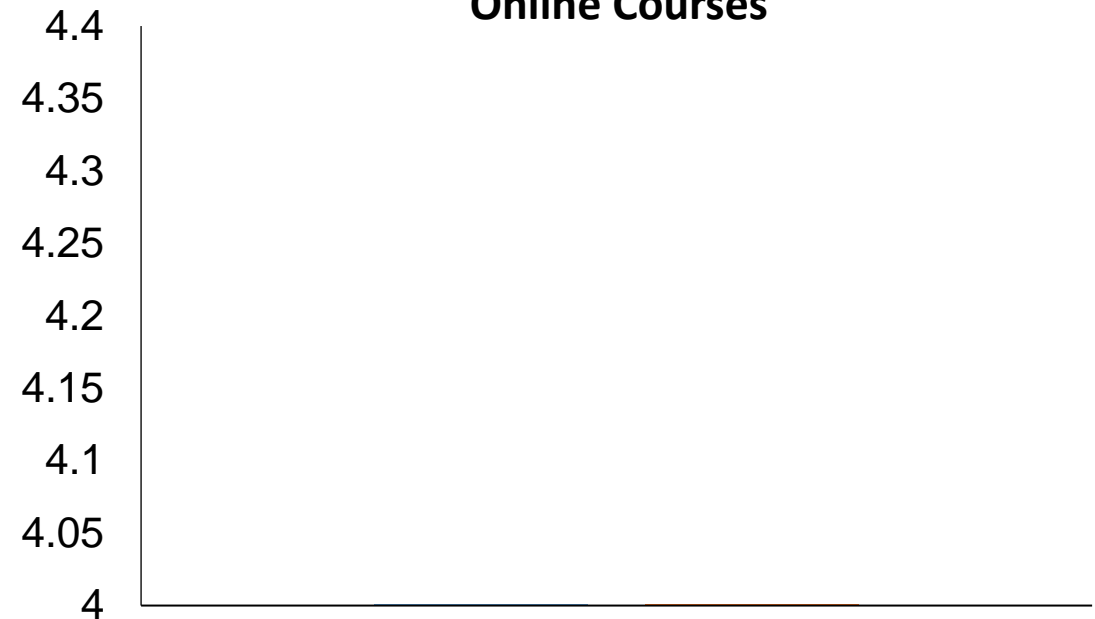
# Pre-Post Comparison: Average SET Scores

Spring 2019 to Spring 2020 Comparison:  
In Person Courses



$t(199) = -3.45, p = .001, \text{Cohen's } d = 0.24$   
 $r_{\text{time}} = .48, p < .0001$

Spring 2019 to Spring 2020 Comparison:  
Online Courses



$t(140) = -1.75, p = .09, \text{Cohen's } d = 0.13$   
 $r_{\text{time}} = .60, p < .0001$

# Change in Average SET Scores: In-Person Courses

Simultaneous Regression				
Predictor	B	$\beta$	SE	<i>p</i>
<b>Instructor Gender</b> 0 = Man, 1 = Woman	.022	.025	.065	.730
<b>Instructor Race</b> 0 = White, 1 = Non-White	.013	.012	.076	.863
<b>Instructor Tenure Status</b> 0 = Tenure, 1 = NT	.009	.009	.070	.899
<b>Course Level</b> 0 = UG, 1 = G	-.114	-.116	.071	.109

Simple Effects		
Group 0 <i>Mdiff</i> (SD)	Group 1 <i>Mdiff</i> (SD)	<i>d</i>
0.05 (0.44)	0.12 (0.49)	0.15
0.07 (0.40)	0.12 (0.61)	0.10
0.09 (0.54)	0.09 (0.42)	< .01
0.12 (0.44)	0.01 (0.44)	0.25



# Change in Average SET Scores: Online Courses

Simultaneous Regression					Simple Effects		
Predictor	B	$\beta$	SE	<i>p</i>	Group 0 <i>Mdiff</i> (SD)	Group 1 <i>Mdiff</i> (SD)	<i>d</i>
<b>Instructor Gender</b> 0 = Man, 1 = Woman	.134	.173	.065	.040	-0.01 (0.37)	0.10 (0.37)	0.30
<b>Instructor Race</b> 0 = White, 1 = Non-White	.079	.094	.070	.262	0.03 (0.39)	0.13 (0.30)	0.26
<b>Instructor Tenure Status</b> 0 = Tenure, 1 = NT	-.026	-.027	.083	.752	0.05 (0.34)	0.06 (0.38)	-0.03
<b>Course Level</b> 0 = UG, 1 = G	-.141	-.186	.064	.030	0.12 (0.38)	0.01 (0.36)	0.30

# Study 1: Summary

---

Overall SET scores for courses taught in Spring 2020 (after SET instructions change and pandemic) were significantly higher than those from the previous spring. This increase:

- Is driven by items about the instructor rather than items about the course
- Is bigger for in-person than online courses

Controlling other factors, increases in evaluations in online (but not in-person) courses were bigger for:

- Women, compared to men, instructors
- Undergraduate, compared to graduate, courses

# Overview of Studies

---

## Study 2

### Experimental, Lab Study SET Items

Participants evaluate a fictitious course and complete SETs that use either standard SET items or best practice SET items

**MOVED STUDY ONLINE,  
ABBREVIATED**

# Study 2 Overview and Participants

## Participants

1816 volunteers at the Project Implicit research website (implicit.harvard.edu) who have taken at least one college course and passed the manipulation check

## Design

2 (Perceived Instructor Gender: Man, Woman) X 2 (SET Items: Standard, Revised)

## Method/Materials

Review of fictitious instructor and course materials

## Gender

Women	71.4%
Men	27.9%

## Race

White	73.1%
Non-White	26.5%

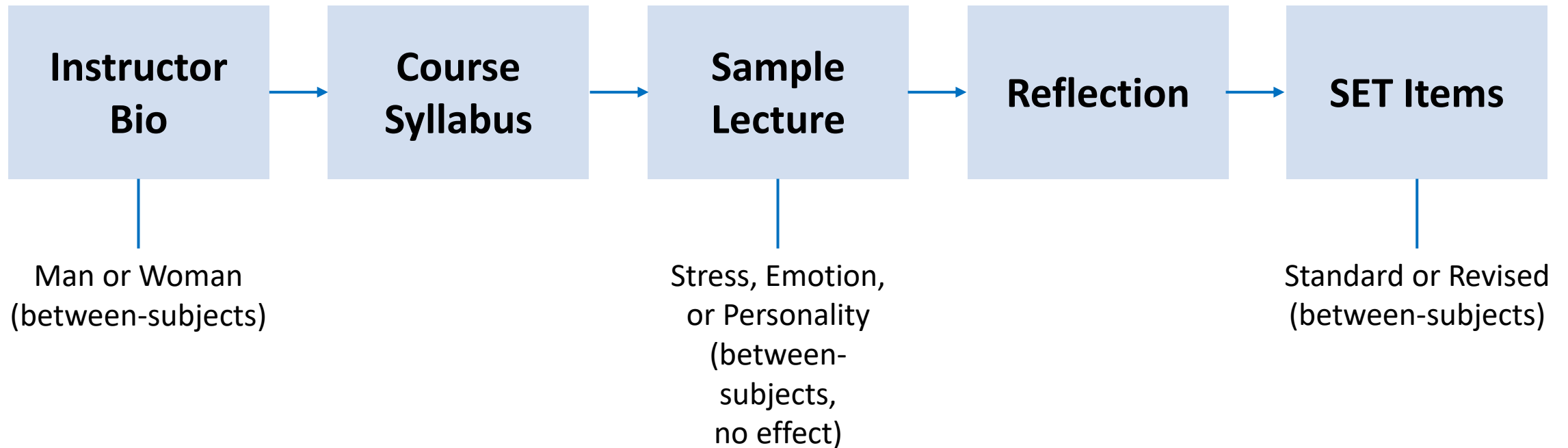
## Education

1-4 classes	23%
5-8 classes	77%
9-12 classes	
13-16 classes	63%
17+	37%

**Politics (1-7 = Very Liberal )** 4.50 (1.62)

**Age** 27.03 (11.49)

# Study 2 Method



## Instructor Biography

Dr. Coleman has a B.S. and Ph.D. in Psychology from the University of Iowa. His research focuses on personality assessment across the lifespan. Dr. Coleman has been a faculty member at Washington State University for the past 12 years. Outside of work, his hobbies include camping, reading, and playing board games. He is also an avid volleyball player

Dr. Coleman wants his students to know the following:

*“I am really looking forward to the experience and challenge of teaching and learning with you. Introductory psychology is a more challenging course than most students expect because it covers such a large amount of information. But due to its nature it is easily applied to our own life and the lives of those around us. Let’s work together this semester to ensure a positive outcome.”*

Continue

## Syllabus

In the next section you will read the syllabus for Dr. Coleman's course. A syllabus describes all aspects of the course: goals, content, structure, assignments, grading, due dates, etc. Students can use the syllabus as a guide to the kind of teaching and learning they can expect from the course.

The syllabus you will see is three pages long. Each page will be presented separately. Once you move forward from a page, you cannot go back. Please pay close attention to the material presented--and how it's presented--so that you can make a judgment about the course and instructor.

[Continue](#)

# PSYCHOLOGY 105

Section 8 | T/Th 10:35-11:50 | Todd 125 | 3 credits | SSCI

## CONTACT



Email: xxxxxxx@wsu.edu  
Office: Johnson Tower 217  
Office hours: T 2-3; W 10-11

## COURSE OVERVIEW

Welcome to the exciting discipline of psychology! As a general introduction to psychology, this course will provide a broad overview of the basic terms, principles, and theories related to the scientific study of behavior and mental processes. A solid understanding of research methodology and the biological basis of behavior will be emphasized. Psychology is fundamentally a science and material will be presented and evaluated from a rigorous scientific perspective. You should come away with a deeper understanding of psychology in daily life.

## CLASS EXPECTATIONS

As a student, there are a number of actions that can both increase your learning in the course and enable you to do well. These are my expectations of you:

- 1) Come to class prepared by reading the textbook.
- 2) Always be present in class, both physically and mentally.
- 3) Be an active participant in the course.
- 4) Be respectful to both your classmates and me.
- 5) Do your own work.

Continue



# What Causes Personality?

- ❖ Genetics: Personality is not determined by any single gene, but rather by the actions of many genes working together
- ❖ Learning History: Bundles of habits acquired by classical and operant conditioning
- ❖ Early Life Experiences: Trauma or care can influence personality later in life

Continue

## Reflection

1 out of 11

In the next part of the study we will ask you to evaluate the course you just learned about. You will respond to the type of questions that are typically asked on end-of-semester evaluations. Before moving on, take a moment to think about Dr. Coleman's course and form a general impression based on the course materials you saw.

Some of the questions might be about aspects to the course that you did not directly learn about. Do your best to imagine what the course would be like based on the materials you saw

Continue

Standard Item	Revised Item

<b>Standard Item</b>	<b>Revised Item</b>
The instructor was enthusiastic about the course	The course fostered learning through generating interest in the subject matter.

<b>Standard Item</b>	<b>Revised Item</b>
The instructor was enthusiastic about the course	The course fostered learning through generating interest in the subject matter.
The instructor explained material clearly and in a way that enhanced my understanding.	The course material was explained clearly and in a way that enhanced my understanding.

<b>Standard Item</b>	<b>Revised Item</b>
The instructor was enthusiastic about the course	The course fostered learning through generating interest in the subject matter.
The instructor explained material clearly and in a way that enhanced my understanding.	The course material was explained clearly and in a way that enhanced my understanding.
The instructor maintained clear standards for response and availability (e.g. turnaround time for email, office hours, etc.)	There were clear standards in this course for availability (e.g., how to get in touch with the instructor) and response time (e.g., turnaround time for email).

Standard Item	Revised Item
The instructor was enthusiastic about the course	The course fostered learning through generating interest in the subject matter.
The instructor explained material clearly and in a way that enhanced my understanding.	The course material was explained clearly and in a way that enhanced my understanding.
The instructor maintained clear standards for response and availability (e.g. turnaround time for email, office hours, etc.)	There were clear standards in this course for availability (e.g., how to get in touch with the instructor) and response time (e.g., turnaround time for email).
The instructor fostered a positive learning environment that engaged students.	The course had an engaging, positive learning environment.

Standard Item	Revised Item
The instructor was enthusiastic about the course	The course fostered learning through generating interest in the subject matter.
The instructor explained material clearly and in a way that enhanced my understanding.	The course material was explained clearly and in a way that enhanced my understanding.
The instructor maintained clear standards for response and availability (e.g. turnaround time for email, office hours, etc.)	There were clear standards in this course for availability (e.g., how to get in touch with the instructor) and response time (e.g., turnaround time for email).
The instructor fostered a positive learning environment that engaged students.	The course had an engaging, positive learning environment.
The instructor provided prompt and meaningful feedback on my work and performance in the course.	The performance feedback provided in this course was useful and effective.

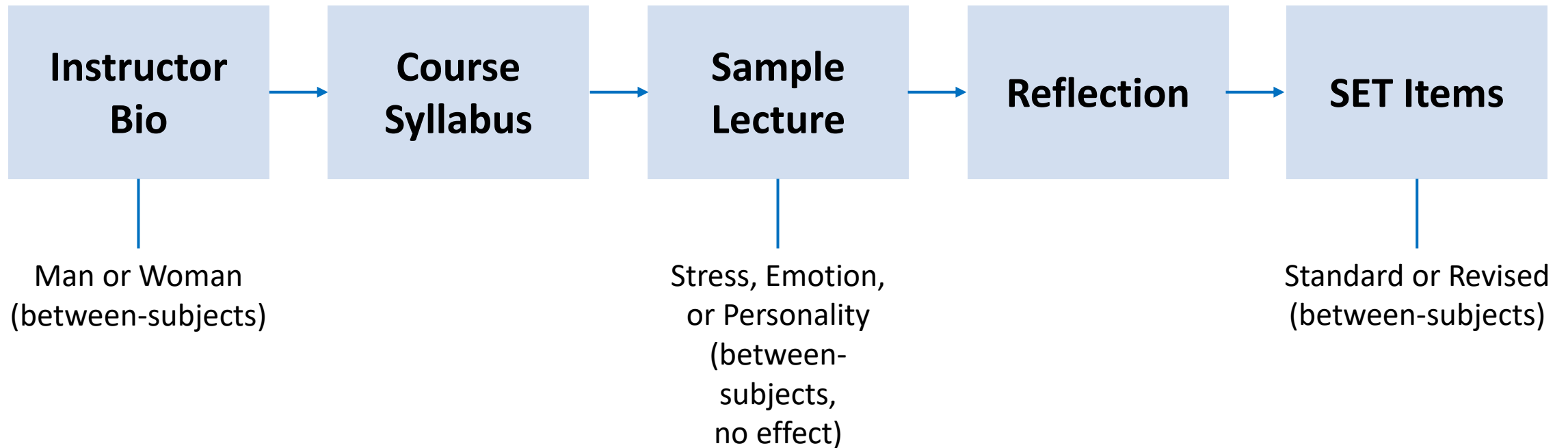


<b>Standard Item</b>	<b>Revised Item</b>
The instructor was enthusiastic about the course	The course fostered learning through generating interest in the subject matter.
The instructor explained material clearly and in a way that enhanced my understanding.	The course material was explained clearly and in a way that enhanced my understanding.
The instructor maintained clear standards for response and availability (e.g. turnaround time for email, office hours, etc.)	There were clear standards in this course for availability (e.g., how to get in touch with the instructor) and response time (e.g., turnaround time for email).
The instructor fostered a positive learning environment that engaged students.	The course had an engaging, positive learning environment.
The instructor provided prompt and meaningful feedback on my work and performance in the course.	The performance feedback provided in this course was useful and effective.
The instructor was instrumental to my learning in the course.	The course fostered learning through high-quality instruction.

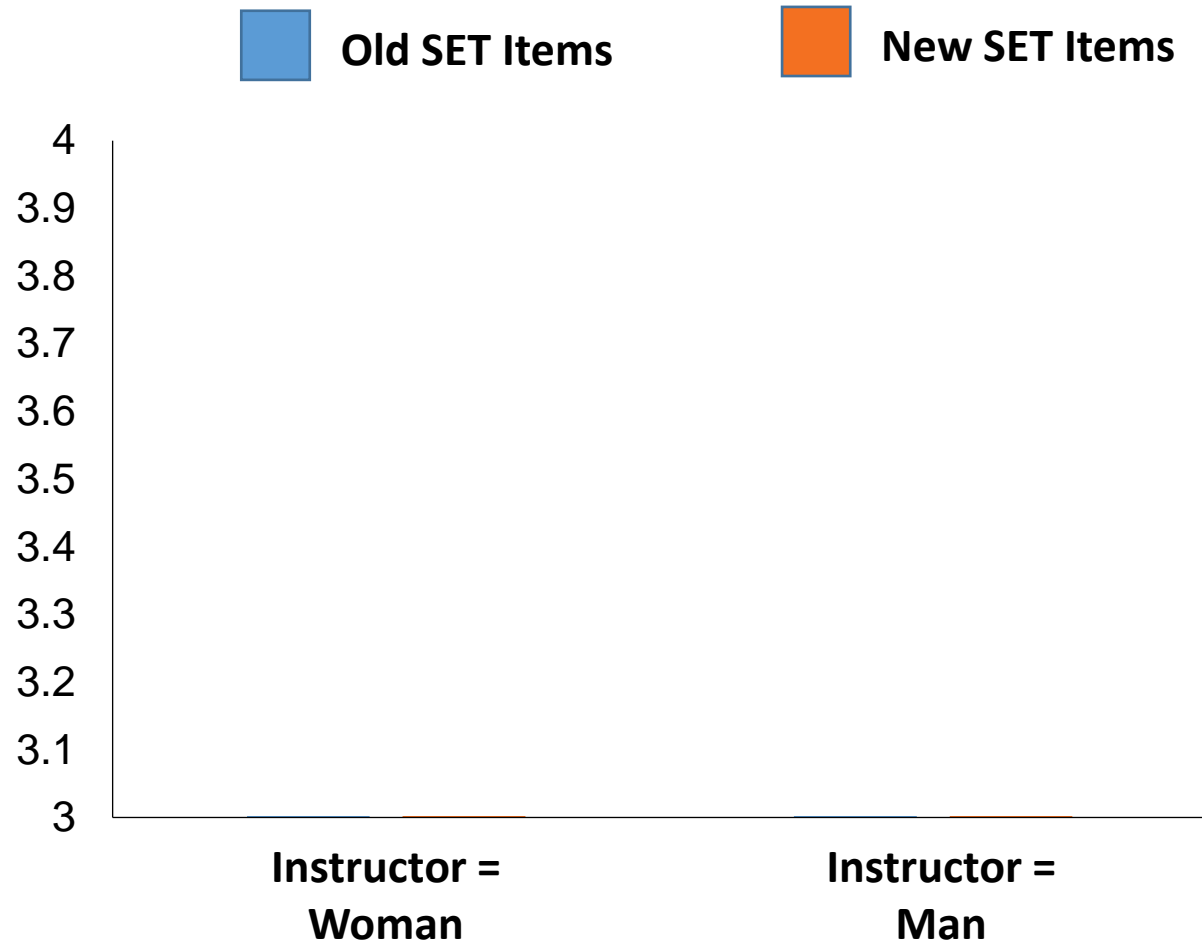
Standard Item	Revised Item
The instructor was enthusiastic about the course	The course fostered learning through generating interest in the subject matter.
The instructor explained material clearly and in a way that enhanced my understanding.	The course material was explained clearly and in a way that enhanced my understanding.
The instructor maintained clear standards for response and availability (e.g. turnaround time for email, office hours, etc.)	There were clear standards in this course for availability (e.g., how to get in touch with the instructor) and response time (e.g., turnaround time for email).
The instructor fostered a positive learning environment that engaged students.	The course had an engaging, positive learning environment.
The instructor provided prompt and meaningful feedback on my work and performance in the course.	The performance feedback provided in this course was useful and effective.
The instructor was instrumental to my learning in the course.	The course fostered learning through high-quality instruction.
The course fostered regular interaction between student and instructor.	The course fostered learning through interaction and engagement.

Standard Item	Revised Item
The instructor was enthusiastic about the course	The course fostered learning through generating interest in the subject matter.
The instructor explained material clearly and in a way that enhanced my understanding.	The course material was explained clearly and in a way that enhanced my understanding.
The instructor maintained clear standards for response and availability (e.g. turnaround time for email, office hours, etc.)	There were clear standards in this course for availability (e.g., how to get in touch with the instructor) and response time (e.g., turnaround time for email).
The instructor fostered a positive learning environment that engaged students.	The course had an engaging, positive learning environment.
The instructor provided prompt and meaningful feedback on my work and performance in the course.	The performance feedback provided in this course was useful and effective.
The instructor was instrumental to my learning in the course.	The course fostered learning through high-quality instruction.
The course fostered regular interaction between student and instructor.	The course fostered learning through interaction and engagement.
The course content (e.g., readings, activities, assignments) was relevant & useful.	No change
Course activities and assignments improved my ability to analyze, solve problems, and/or think critically.	
Overall, this course was a valuable educational experience.	

# Study 2 Method



# Study 2 Results



Predictor	B	$\beta$	SE	<i>p</i>
SET Item Type	-.090	-.081	.042	.020
Instructor Gender	-.084	-.066	.043	.058
Gender*SET Type	.010	.007	.061	.871
Participant Age	-.012	-.046	.001	.054
Gender	.018	.013	.001	.587
Race	.076	.046	.040	.055
Politics	-.013	-.033	.010	.176
Education	.001	.005	.004	.823

## Study 2: Summary and Future Direction

---

Controlling for a number of demographic factors, participants evaluated a novel, fictitious course more favorably when using the standard, current SET items compared to revised, "improved" SET items

### Study 3 (Future)

#### **Experimental, Field Study SET Items**

Using what we learn in Study 2, UF students will be randomly assigned to complete actual SETs that use either standard SET item or best practice SET items

More research is necessary before implementing revised instructions:

- UF student sample
- More in-depth, realistic course materials

UF

Thank you! Contact: [ratliff@ufl.edu](mailto:ratliff@ufl.edu)