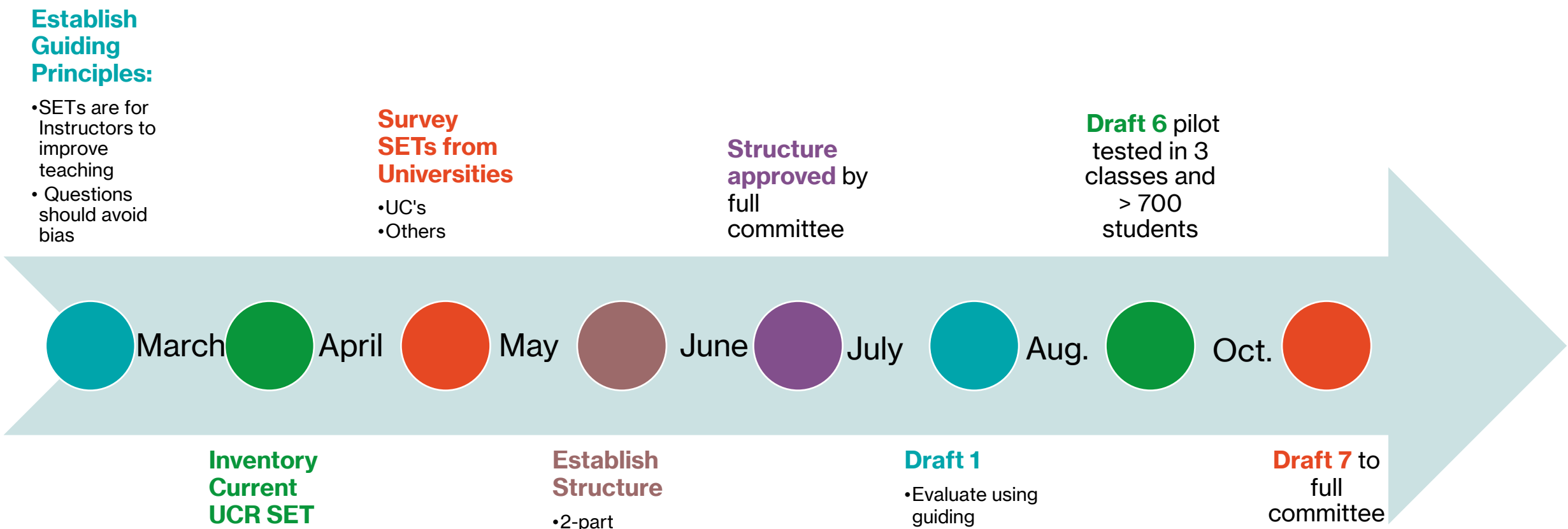
Appendix 2

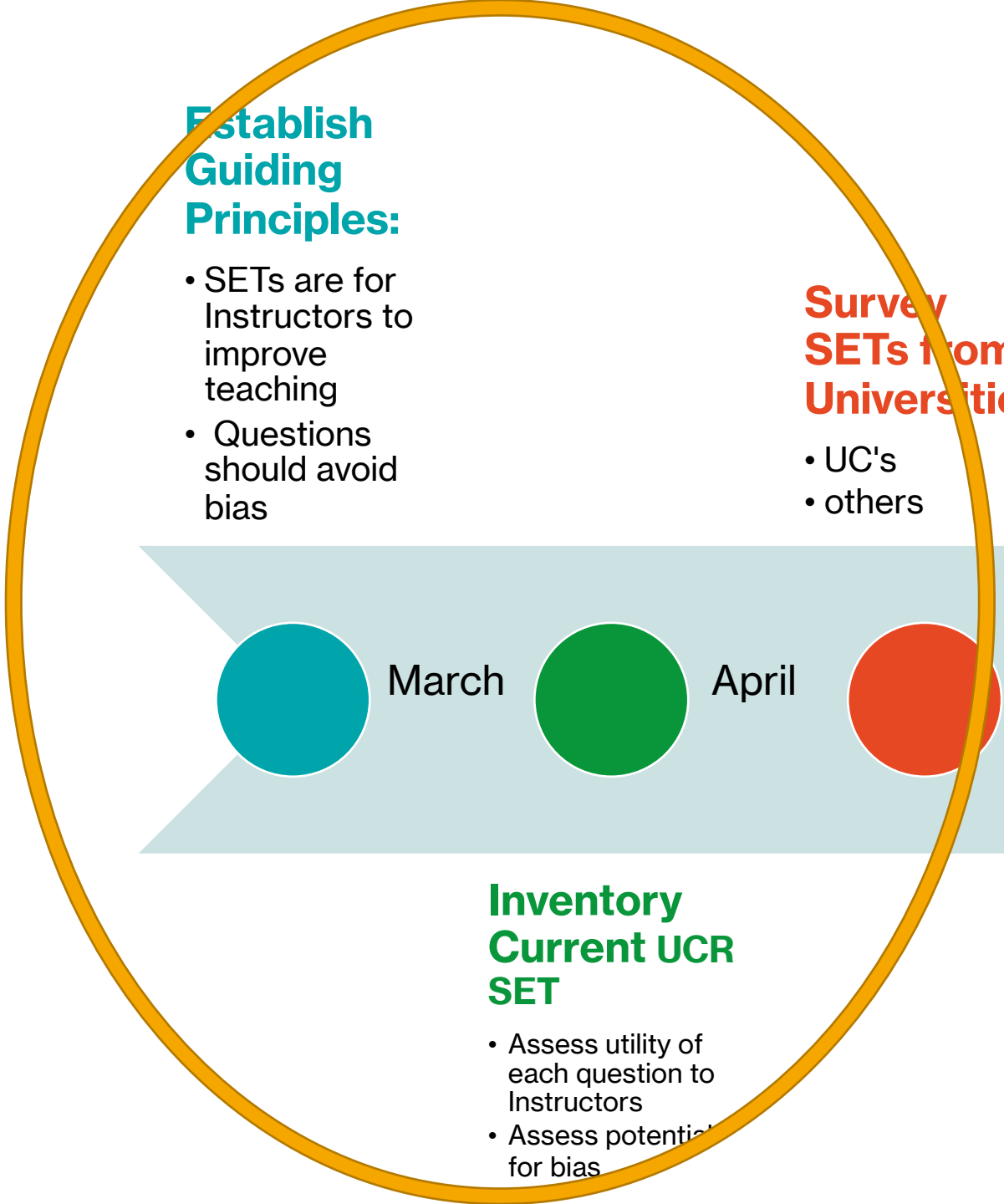
Report on the Creation of the New Instrument for SETs

Prepared by the Question Wording Subcommittee of the
Teaching Evaluation Implementation Committee (TEIC)

Co-chaired by: Annie S. Ditta & Goldberry Long

Members: Ahmed Eldawy, Hillary Jenks, Omar Safie, Patriccia Ordonez-Kim, & Batool Abdaljawad


Establish Guiding Principles:
•SETs are for Instructors to improve teaching
• Questions should avoid bias
Survey SETs from Universities
•UC's
•Others
Structure approved by full committee
Draft 6 pilot tested in 3 classes and > 700 students
March
April
May
June
July
Aug.
Oct.
Inventory Current UCR SET
•Assess utility of each question to Instructors
•Assess potential for bias
Establish Structure
•2-part deployment
•4 domains based on surveyed SETs and literature
•2-part questions
•open comments in each domain
Draft 1
•Evaluate using guiding principles
•Rewrite where necessary
•Draft 2, 3 - repeat process
•Draft 4 to full committee
Draft 5, 6
Draft 7 to full committee



Establish Guiding Principles:

- SETs are for Instructors to improve teaching
- Questions should avoid bias

Survey SETs from Universities

- UC's
- others

Structure approved by full committee

March

April

May

June

July

Inventory Current UCR SET

- Assess utility of each question to Instructors
- Assess potential for bias

Establish Structure

- 2-part deployment
- 4 domains based on surveyed SETs and literature

Draft 1

- Evaluate u guiding principles
- Rewrite w necessary
- **Draft 2, 3** repeat

Guiding Principle 1:

SETs are for Instructors to improve their teaching

1

SETs should **give feedback** on instructional practices

2

This feedback should be **usable for improving teaching**

3

Questions that don't meet these 2 criteria should not be included

Guiding Principle 2:

Questions should avoid bias by:

1

Focusing **on instructor behavior** and **curriculum,** not student feelings

2

Relying on best practices for evaluation by addressing items that are **specific, observable, and achievable**
(Stiggins, 1987)

3

Avoiding vague, subjective, personality-based items that invite emotions or value judgments, which are prone to bias*

*Examples of bias found in Kreitzer & Sweet-Cushman (2022)

Some examples of "vague, subjective, personality-based items"

Enthusiastic

- Some instructor's enthusiasm **may not be observable** for cultural or personality reasons
- Evidence suggests **enthusiasm is not associated** with increased learning

Prepared

- Extensive preparation **may not be observable** by students, especially in active-learning modes

Caring

- Making a student feel cared for **might not be achievable**, especially if they aren't doing well in the class
- Caring is **not always observable**

Organized

- **Not specific**; includes many behaviors and actions,
- Hard to know what behavior to change.
- e.g., if student says an instructor is disorganized, that might mean:
  - Poor communication or responsiveness
  - Lacking course foundations
  - Unclear deadlines and course schedule
  - Didn't return assessments in a timely manner

Some examples of specific, observable, achievable items

Did the instructor provide. . .

- Criteria for grading
- Due dates
- Modes of communication
- Opportunities to participate in classroom discussion

These criteria are described in Graham et al. (2001)

Establish Guiding Principles:

- SETs are for Instructors to improve teaching
- Questions should avoid bias

Survey SETs from Universities

- UC's
- others

Structure approved by full committee

Draft 6 pilot test in > 700 students

March | April | May | June | July | Aug. | Oct.

Inventory Current UCR SE

- Assess utility of each question to Instructors
- Assess potential for bias

Establish Structure

- 2-part deployment
- 4 domains based on surveyed SETs and literature
- 2-part questions
- open comments in each domain

Draft 1

- Evaluate using guiding principles
- Rewrite where necessary
- **Draft 2, 3 -** repeat process
- **Draft 4** to full committee

Draft 5, 6

Timeline of Evaluations

Part 1
Administered in Weeks 2-3

A) Course Foundations 1

- **Course structure** that should be present at start of a class -- syllabus, grading scheme, etc.
- **Students asked** if items are present
- **Gives instructor a chance** to update/improve course
- **Shows students** that their SETs feedback is used

Part 2
Administered in Weeks 8-9

A) Course Foundations 2
results replace Part I results

B) Class Experiences

C) Assessment

D) Learning Support

- **Same list** as part 1 except 2 Q's:
- Did instructor inform you of any changes to this list in a timely manner?
- What, if anything, on this list would you suggest the instructor change for learning of future students?

For domains B,C,D, **each question** follows this 2-part format:

1. **Instructor's Responsibility for item**:.
   "Did instructor provide X?"
   Absolutely/Somewhat/Not Really

2. **Student experience** of course item:
   "Was this useful to your learning?"
   Not Useful/Useful/Very Useful

At the end of each domain, students are asked for **Advice**: (open-ended) about that domain.

Why these 4 domains?

- **Inventory of previous SETs** revealed 4 themes that align with literature* (Graham et al., 2022; Patrick & Smart, 1998; Tinto, 2012)
- **Report to instructor** will group data in 4 categories of best practice
- **Allows instructors** to target specific areas for improvement

A) Course Foundations 2
results replace Part I results

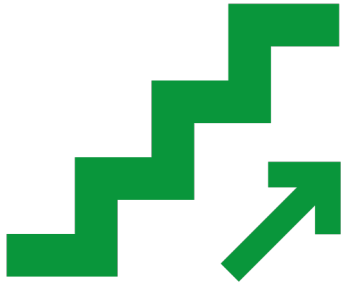
B) Class Experiences

C) Assessment

D) Learning Support

*Except in cases where such items may introduce bias due to not being observable (e.g., “organization” as previously described)

Why are Domain A items called “Course Foundations?”



Items involve organization, materials, clarity of expectations, goals, and modes of communication, all needed at the start of class

Evidence suggests that providing these early optimizes student learning (Tinto, 2012)

Are you telling me what to do in my class?

- Good evaluation uses clearly defined standards
- This SET is meant to evaluate good teaching practice
- All items are best teaching practices that evidence suggests should be present for students to achieve optimum learning (Lang, 2021)
- Relying on concrete, evidence-based criteria gives the best chance of avoiding bias
- All items are specific, observable, measurable, and achievable; you choose whether to include them

Part 1
Administered in Weeks 2-3

A) Course Foundations 1

- **Course structure** that should be present at start of a class -- syllabus, grading scheme, etc.
- **Students asked** if items are present
- **Gives instructor a chance** to update/improve course.
- **Shows students** that their SETs feedback is used

Why administer Part 1 early?

- To show students that their feedback is valuable and contributes to observable change in teaching (Chen & Hoshower, 2003)
- To give instructors the best chance to demonstrate their excellence in teaching by responding to feedback

<u>Part 1: Course Foundations Check</u>: *All of the following provide a framework for learning in this course. These items may be found in the syllabus, on Canvas, and/or in handouts. Are they present in at least one of these places?*

1. A clear description of what you should be able to do or know by the end of the course (learning outcomes). Yes/No
2. A clear description of the grading system for this class Yes/No
3. Information on how to ask for help (for example: office hours, email, Zoom appointments, etc.) Yes/No
4. Information on how to ask for accommodations/support from the university (for example: Contact information for SDRC, Title IX office, CAPS, ARC, etc.) Yes/No
5. A list of course topics Yes/No
6. Assignment due dates Yes/No
7. Guidelines for academic integrity Yes/No
8. Course materials (may include: readings, software, textbooks, recordings, and other resources) Yes/No
9. Were the above items compiled in an accessible location (e.g., a document, a Canvas page, or other format)? Yes/No
10. Are you enrolled in a lab for this course? IF NO, skip 11. Yes/No
11. IF YES: have you received instructions for taking appropriate safety measures in lab settings? Yes/No

Why a yes/no answer for Course Foundations?

- These are concrete items that are either present or they are not
- Students can’t provide further feedback this early in the quarter

No

Yes

Instructor will receive a Part 1 report early in the quarter

- Percentage of items present
- Instructor is free to change course items in response to feedback (i.e., address omissions of critical course components early)
- Part 2 results will overwrite Part 1
  - Allows the feedback to not be punitive by becoming a part of the official evaluation

Part 2: Administered week 8-9

A)
Course Foundations 2
results replace Part I results

B)
Class Experiences

C)
Assessment

D)
Learning Support

For domains B,C,D, **each question** follows this 2-part format:

1. **Instructor's Responsibility for item**:.
   "Did instructor provide X?"
   Absolutely/Somewhat/Not Really

2. **Student experience** of course item:
   "Was this useful to your learning?"
   Not Useful/Useful/Very Useful

At the end of each domain, students are asked for **Constructive Advice** (open-ended) about that domain.

Why the two-part question?

- Research suggests students aren't very good at evaluating their own learning (Carpenter et al., 2021)
- We are separating the *presence of the evidence-based practice* from the *students' perception* of whether it was useful
  - Allows more nuance when interpreting and using the results

Why are you using a 3-point instead of a 5-point Likert scale?

- “working with fewer scale values increases the reliability of an instrument” (Spooren et al., 2007)

- To use the SET as a tool to improve teaching**, instructors need to know if an item was useful** for students’ learning
- Instructor will keep an item that is ***useful*** or ***extremely useful***; a more fine-grained numeric scale won’t change that choice
- If something is ***not useful***, the instructor will consider changing it; the instructor doesn’t need to know if it is ranked 0, 1 or 2 of ”not useful”
- The most important information is ***what percentage of students*** said an item was not useful

Why the informal language in the response options?

- Familiar language allows students to intuitively understand questions
- Therefore, familiar language is most likely to capture accurate responses

Why an open-ended question at the end of each domain?

- Linking open comments to specific, observable teaching practices limits chances of ad hominem attacks
- Comments can be grouped by domain, allowing for instructor to identify key areas that need revision

Why aren’t you asking students how hard they worked?

Instrument is for instructors	• This instrument assesses teaching, not student behavior • Students are not very good at assessing their own level of work, especially if they find the class challenging • Instructors are not able to change the level of work students do, which means they can’t act on the information
Instrument should avoid bias	• While a question about how hard they worked may seem to debias students, it might actually increase bias by engaging emotions like frustration and anger
Questions should be specific and observable	• Students’ perception of work level is not observable or specific

References

Calvin D. Smith & Chi Baik (2021) High-impact teaching practices in higher education: a best evidence review, Studies in Higher Education, 46:8, 1696-1713, DOI: 10.1080/03075079.2019.1698539 Hattie, J. 2008. *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. London: Routledge

CHEN, Y &. HOSHOWER, L(2003) Student Evaluation of Teaching Effectiveness: An assessment of student perception and motivation, Assessment & Evaluation in Higher Education, 28:1, 71-88, DOI: 10.1080/02602930301683

Graham, C., Cagiltay, K., Lim, B. R., Craner, J., & Duffy, T. M. (2001). Seven principles of effective teaching: A practical lens for evaluating online courses. *The Technology Source*, *30*(5), 50.

Hattie, J. 2008. *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. London: Routledge

Kreitzer, R. J., & Sweet-Cushman, J. (2022). Evaluating student evaluations of teaching: A review of measurement and equity bias in SETs and recommendations for ethical reform. *Journal of Academic Ethics*, *20*(1), 73-84.

Lang, J. M. (2021). *Small teaching: Everyday lessons from the science of learning*. John Wiley & Sons.

Patrick, J., & Smart, R. M. (1998). An empirical evaluation of teacher effectiveness: The emergence of three critical factors. *Assessment & Evaluation in Higher Education*, *23*(2), 165-178.

Shana K. Carpenter, Amber E. Witherby, Sarah K. Tauber, On Students' (Mis)judgments of Learning and Teaching Effectiveness, Journal of Applied Research in Memory and Cognition, Volume 9, Issue 2, 2020, pp 137-151 https://doi.org/10.1016/j.jarmac.2019.12.009

Spooren, P. Mortelmans, D & Denekens, J. (2007) Student evaluation of teaching quality in higher education: development of an instrument based on 10 Likert-scales, Assessment & Evaluation in Higher Education, 32:6, 667-679, DOI: 10.1080/02602930601117191

Stiggins, R. J. (1987). Design and development of performance assessments. *Educational measurement: Issues and practice*, *6*(3), 33-42.

Tinto, V. (2012). Enhancing student success: Taking the classroom success seriously. *Student Success*, *3*(1), 1.