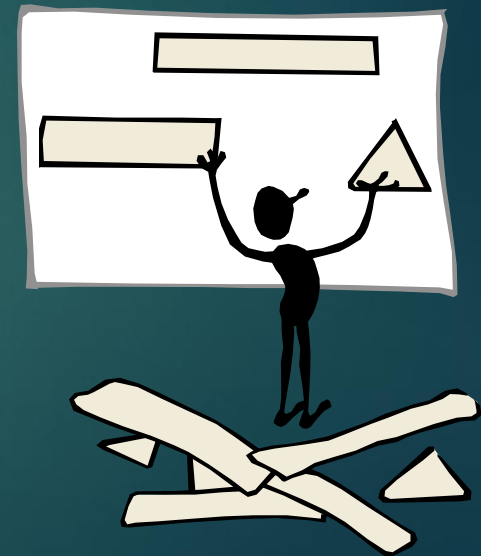


Examination of the Student Evaluation of Teaching Form 2015

DALE PIETRZAK, ED.D., LPC-MH, CCMHC
(DIRECTOR, INSTITUTIONAL EFFECTIVENESS AND ACCREDITATION)
UNIVERSITY OF IDAHO



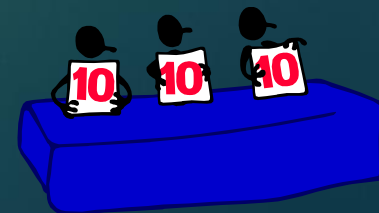
There is no Magic



As with nearly all assessments, there is no perfect model, so we are looking for a functional one.

General Overview

- ▶ The SET form is not intended to be used in isolation, but is required as part of a process to examine both teaching and course structure.
- ▶ There is an expectation that each College has an established teaching and course review process, of which this is to be one part.
 - ▶ IEA recommends that this process include other evaluation procedures for teaching and course assessment to supplement and provide context for the results of this tool.
- ▶ For example, the tool is part of the process which may be used at the university level in the context of DFWI reports (included in these are course GPA and % A grades) and other academic measures. The tool is not intended to be reviewed in isolation.
- ▶ This contextual use assists in providing a checks-and-balances process to ensure rigor in coursework as well as examination of teacher and course elements.



Personal Values



- ▶ There are few areas in higher education that have the emotional loading that a student evaluation of teaching carries. Nearly every faculty member believes himself/herself to be an excellent instructor and with well-run and interesting courses. There are few “average” self-ratings here.
 - ▶ Practically every instructor can locate a student in a course to provide support for them.
 - ▶ Those wishing to find support for the negative input can generally find a student to support that view as well.
 - ▶ There has been enough research in this area (generally very poorly designed with few controls for extraneous variables) that one can find support for nearly any position. Much like student support, one can also find the support of an opposing view in the literature as well.

What now?

- ▶ A general overview of the research in the area
 - ▶ Pietrzak, D. Duncan, K. & Korcuska, J. (2008). Counseling students' decision making regarding faculty teaching effectiveness: A Conjoint Analysis. *Counselor Education and Supervision*, 48(2), 114-132.
- ▶ Selected well designed studies or interesting studies
 - ▶ Al-Busaidi, Aldhafri, S. & Buyukyavuz, O. (2016). Effective university Instructors as perceived by Turkish and Omani university students. *Sage Open*, July-Sept., 1-8.
 - ▶ Marsh, H. & Roche, L. (2000). Effects of grading leniency and low workload on students' evaluations of teaching: Popular myth, bias, validity or innocent bystanders. *Journal of Educational Psychology*, 92(1), 202-228.
 - ▶ Overall, J. & Marsh, H. (1980). Students' evaluations of instruction: A longitudinal study of their stability. *Journal of Educational Psychology*, 72(3), 321-325.
 - ▶ Remedios, R. & Lieberman, D. (2008) I liked your course because you taught me well: The influence of grades, workload, expectations and goals on students evaluations of teaching. *British Educational Research Journal*, 34(1) 91-1125.



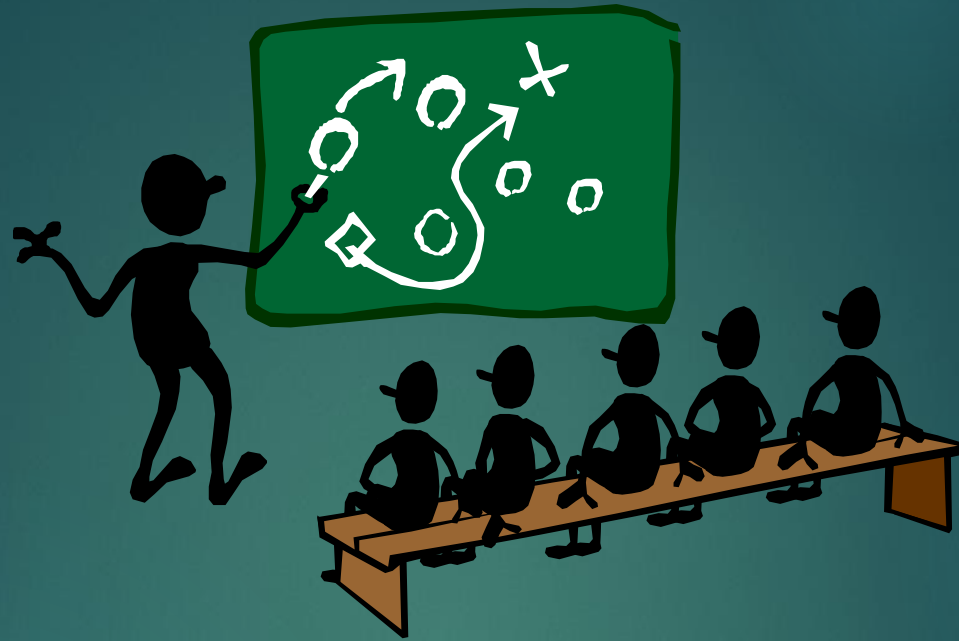
What now? continued

- ▶ Maurer, T. (2006). Cognitive Dissonance or Revenge? Student Grades and Course Evaluations. *Teaching of Psychology*, 33(3), 176-179.
- ▶ Spooren, P. & Mortelmans, D. (2006). Teacher professionalism and student evaluation of teaching: will better teachers receive higher ratings and will better students give higher ratings? *Educational Studies*, 32(4) 201-214.
- ▶ Addison, W., Best, J. & Warrington, J. (2006). Students' perceptions of course difficulty and their ratings of the instructor. *College Student Journal*, 40(2).
- ▶ Summary – It is evident that grades and other elements have some relationship to instructor evaluations. It is believed there *should* be some relationship. However, the current data seem to suggest these forms pick up largely a “classroom environment” factor. This is not really a new development. No assessment measures one thing. The question really is, “Does the effect of these elements negate the utility of these types of tools?” It is also apparent that course level, student characteristics, course content and subject matter each impact performance. These need to be controlled in studies.

So Now What?

▶ Summary:

- ▶ It is evident that grades and other elements have some relationship to instructor evaluations. It is believed there *should* be some relationship to grades. This is not really new in the assessment world. The question really is, “Does the **effect** of these elements negate the **utility** of these types of tools?” It is essentially a Convergent and Discriminant Validity study process.
- ▶ ***However, the data to date seem to suggest these forms pick up a “classroom environment” factor.***



Relationship to Grades

SHOULD GRADES BE RELATED?

Questions & Sample

▶ **Core Questions**

- ▶ How related are grades to Student Evaluation?
- ▶ How do Student Evaluation scores relate to other factors?

▶ **Sample**

- ▶ All available Student Evaluation of Teaching scores from Fall 2012 to Summer 2015
 - ▶ A sample of 16,371 class average scores
 - ▶ This represents 64,845 student responses

▶ **Guide to Correlations:**

- ▶ <https://statistics.laerd.com/statistical-guides/pearson-correlation-coefficient-statistical-guide.php>

Relationship to Grades: All Courses

| Pearson Correlation | | |
|----------------------------|--|--|
| | Overall, how would you rate the instructor's performance in teaching this course? | Overall, how would you rate the quality of this course? |
| Course GPA | 0.28 | 0.28 |
| Percent A | 0.27 | 0.27 |
| Percent W | -0.15 | -0.16 |
| Percent C or Better | 0.22 | 0.23 |

N=12,187 UG and GRAD courses

Relationship to Grades: Undergraduate Courses

| Pearson Correlation | | |
|--------------------------------|--|--|
| | Overall, how would you rate the instructor's performance in teaching this course? | Overall, how would you rate the quality of this course? |
| Course GPA | 0.28 | 0.29 |
| Percent A | 0.27 | 0.27 |
| Percent W | -0.16 | -0.16 |
| Percent C or Better | 0.22 | 0.24 |
| <i>N=10,136 Courses</i> | | |

Relationship to Grades: Graduate Courses



| Pearson Correlation | | |
|----------------------------|--|--|
| | Overall, how would you rate the instructor's performance in teaching this course? | Overall, how would you rate the quality of this course? |
| Course GPA | 0.15 | 0.12 |
| Percent A | 0.15 | 0.12 |
| Percent W | -0.05 | -0.06 |

N=2,051 Graduate Courses

Relationship to Grades: Course Level 100-200



| Pearson Correlation | | |
|--------------------------------|--|--|
| | Overall, how would you rate the instructor's performance in teaching this course? | Overall, how would you rate the quality of this course? |
| Course GPA | 0.37 | 0.38 |
| Percent A | 0.37 | 0.37 |
| Percent W | -0.22 | -0.21 |
| Percent C or Better | 0.29 | 0.31 |
| N=4,233 100-200 Courses | | |

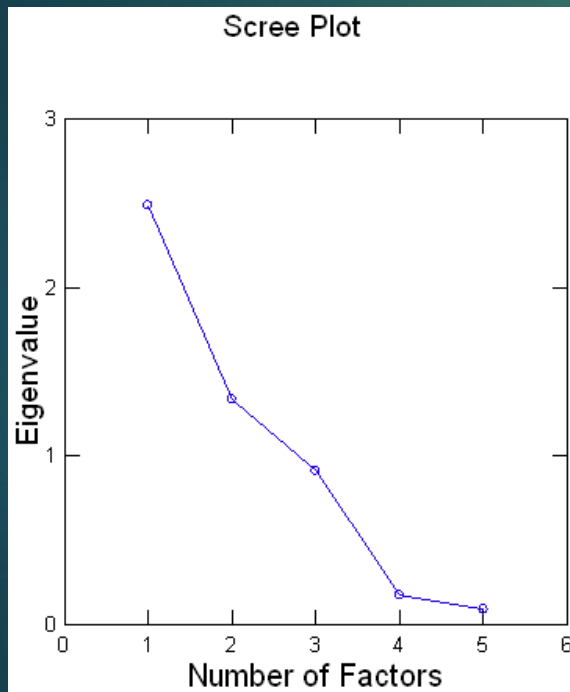
Relationship to Grades: Course Level 300-400



| Pearson Correlation | | |
|--|--|--|
| | Overall, how would you rate the instructor's performance in teaching this course? | Overall, how would you rate the quality of this course? |
| Course GPA | 0.24 | 0.25 |
| Percent A | 0.25 | 0.25 |
| Percent W | -0.11 | -0.11 |
| PCT_SUCC | 0.17 | 0.18 |
| <i>N=10,136 300-400 Courses</i> | | |

Principle Components: Oblique

- ▶ A 2 or possibly 3 factor solution seem to fit the instructor, course and academic performance data.



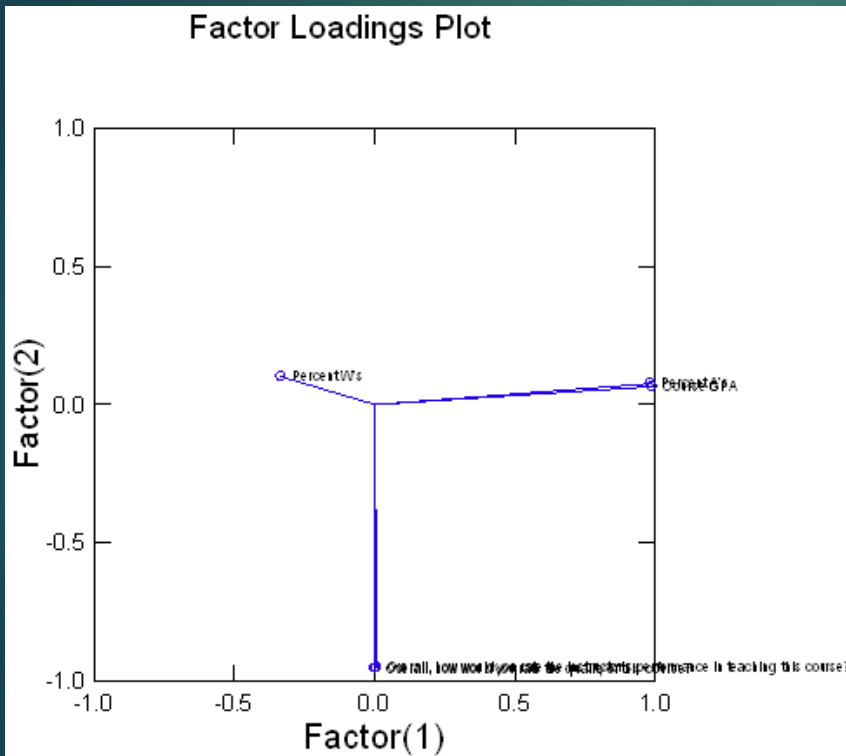
| 2 Factor solution accounting for 76.4% of the variance | | |
|--|-------|-------|
| | F1 | F2 |
| Instructor | 0.01 | -0.95 |
| Course | 0.00 | -0.95 |
| Course GPA | 0.99 | 0.07 |
| Percent A | 0.98 | 0.08 |
| Percent W | -0.33 | 0.10 |

| 3 Factor Solution accounting for 94.7% of the variance | | | |
|--|------|------|-------|
| | F1 | F2 | F3 |
| Instructor | 0.00 | 0.96 | 0.01 |
| Course | 0.00 | 0.96 | -0.01 |
| Course GPA | 0.97 | 0.00 | -0.03 |
| Percent A | 0.98 | 0.00 | 0.03 |
| Percent W | 0.00 | 0.00 | 1.00 |

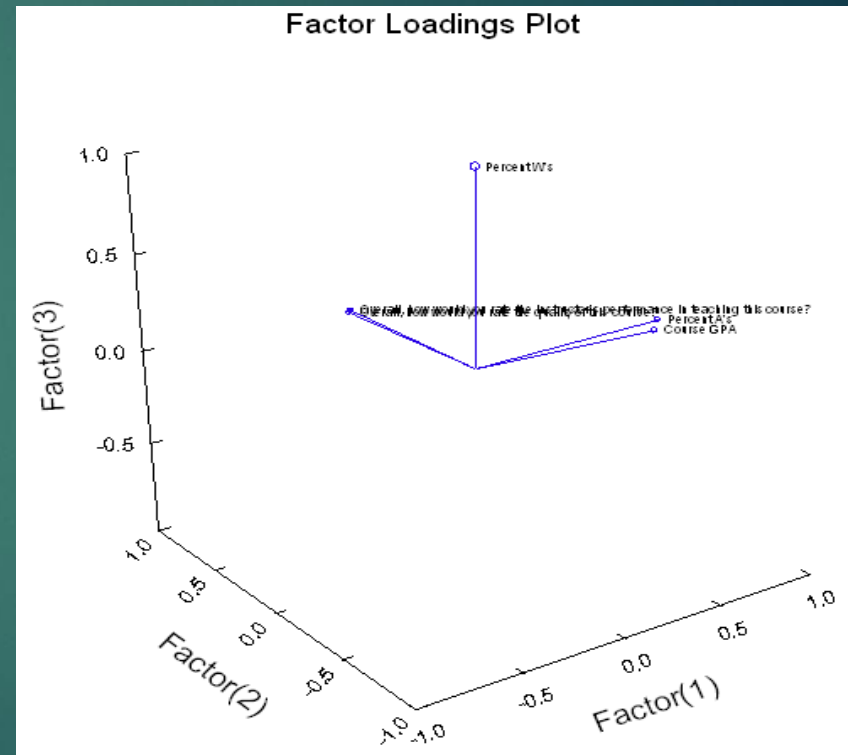
| Correlations Among Oblique Factors | | | |
|------------------------------------|-------------|--------------|------|
| | F1 | F2 | F3 |
| F1 | 1.00 | | |
| F2 | 0.29 | 1.00 | |
| F3 | -0.21 | -0.16 | 1.00 |

PCA Loading Plot

While the text is too small to read, the graphics provide a visual representation of the relationships. The Students Evaluation items are at nearly perfect right angles to academic metrics which suggests little shared variance.



Two Factor Model



Three Factor Model

Observations

- ▶ The current data find a small relationship between grades and the Teacher or Course scores, which is expected.
 - ▶ One would expect that there be a correlation in the .20 to .35 range, based on the research literature to date. It should be neither too high, nor non-existent.
 - ▶ These data are consistent with theory that these forms tap the educational atmosphere of the classroom. In general students in a more positive learning environment benefit.



What Other Elements do the Teaching Scores Relate to at UI?

Instructor Gender and Ethnicity

| Pearson Correlation Matrix | | |
|------------------------------------|---|---|
| | Overall, how would you rate the instructor's performance in teaching this course? | Overall, how would you rate the quality of this course? |
| GENDER (N=14776/14770) | 0.00 | 0.01 |
| MINORITY (N=14776/14770) | -0.03 | -0.04 |
| CITIZEN (N=14776/14770) | -0.03 | -0.03 |

Relationship to other items: Highest (.80 >)

| Overall, how would you rate the <u>INSTRUCTOR</u> ? | <u>Correlation</u> |
|---|--------------------|
| Q101 Clarity of instructor's explanations | 0.92 |
| Q137 Likelihood you would recommend this instructor to others | 0.91 |
| Q129 Instructor's ability to stimulate interest in the course topics | 0.89 |
| Q128 Instructor's ability to teach critical thinking on the course topics | 0.89 |
| Q102 Logical presentation of course material by the instructor | 0.88 |
| Q106 Instructor's ability to convey relevance of course material | 0.88 |
| Q103 Presentation of course material by the instructor | 0.86 |
| Q122 Instructor's skill in interpreting student responses | 0.85 |
| Q120 Helpfulness of answers to student questions | 0.85 |
| Q113 Helpfulness of instructor's lectures in understanding the material | 0.85 |
| Q115 Instructor's concern for quality of teaching | 0.83 |
| Q109 Instructor's use of multimedia presentations | 0.83 |
| Q123 Presentation of material at a level appropriate to the audience | 0.82 |
| Q111 Instructor's use of instructional methods | 0.82 |
| Q118 Instructor's ability to handle unexpected questions | 0.82 |
| Q104 Instructor's use of class time | 0.81 |
| Q139 Instructor's ability to build a sense of community in the classroom. | 0.81 |
| Q125 Instructor's availability for help outside of class | 0.81 |
| Q110 Clarity of instructor's verbal communication | 0.81 |
| Q119 Instructor's encouragement for subject matter | 0.80 |
| Overall, how would you rate the quality of this course? | 0.80 |
| Q105 Instructor's ability to make course material interesting | 0.80 |
| Q136 Instructor's commitment to safety in the laboratory | 0.80 |

Relationship to other items: Lowest

| Overall, how would you rate the <u>INSTRUCTOR</u>? | <u>Correlation</u> |
|--|---------------------------|
| 001 What grade do you expect to receive? | 0.37 |
| 003 How would you rate the quality of your effort in this class? | 0.29 |
| 004 What grade were you working to achieve? | 0.25 |
| 005 How often did you attend class? | 0.22 |
| 002 How often were you fully prepared for class? | 0.22 |

The pattern of correlations with other items (what is high versus what is low) supports the measurement of the educational atmosphere of the classroom as the core construct. These items generally relate to instructor-controlled elements of the course.

Relationship to other items: Highest (.83 >)

| Overall, how would you rate the quality of this <u>COURSE</u> ? | Correlation |
|---|-------------|
| Q664 Emphasis on creativity and original thought in the exams | 0.96 |
| Q337 Course's value in teaching how to evaluate new work in this subject | 0.94 |
| Q395 Course's value in teaching the value of new viewpoints | 0.93 |
| Q363 Course's value in stimulating interest in outside reading | 0.92 |
| Q488 Course's value in developing your ability to use and integrate information from multiple sources. | 0.91 |
| Q541 Relevance of written assignments to course materials | 0.90 |
| Q543 Quality of written assignments in capturing student interest | 0.89 |
| Q396 Course's value in encouraging reconsideration of former attitudes | 0.89 |
| Q340 Course's value in teaching written communication skills in the subject matter | 0.88 |
| Q424 Course's value in gaining an awareness of interests and talents | 0.88 |
| Q332 Course's value in gaining an understanding of the subject matter | 0.87 |
| Q544 Quality of written assignments in stimulating student thought | 0.87 |
| Q489 Course's value in developing your ability to think through real-world issues, explore creative avenues of expression, solve prob.. | 0.87 |
| Q362 Course's value in stimulating enthusiasm for learning the subject matter | 0.87 |
| Q304 Appropriateness of the level at which course material is covered | 0.86 |
| Q492 Course's value in developing your critical thinking skills by learning how to identify and evaluate arguments. | 0.86 |
| Q574 Quality of reading assignments in stimulating student thought | 0.85 |
| Q361 Course's value in stimulating interest in the subject | 0.85 |
| Q338 Course's value in teaching critical thinking skills | 0.85 |

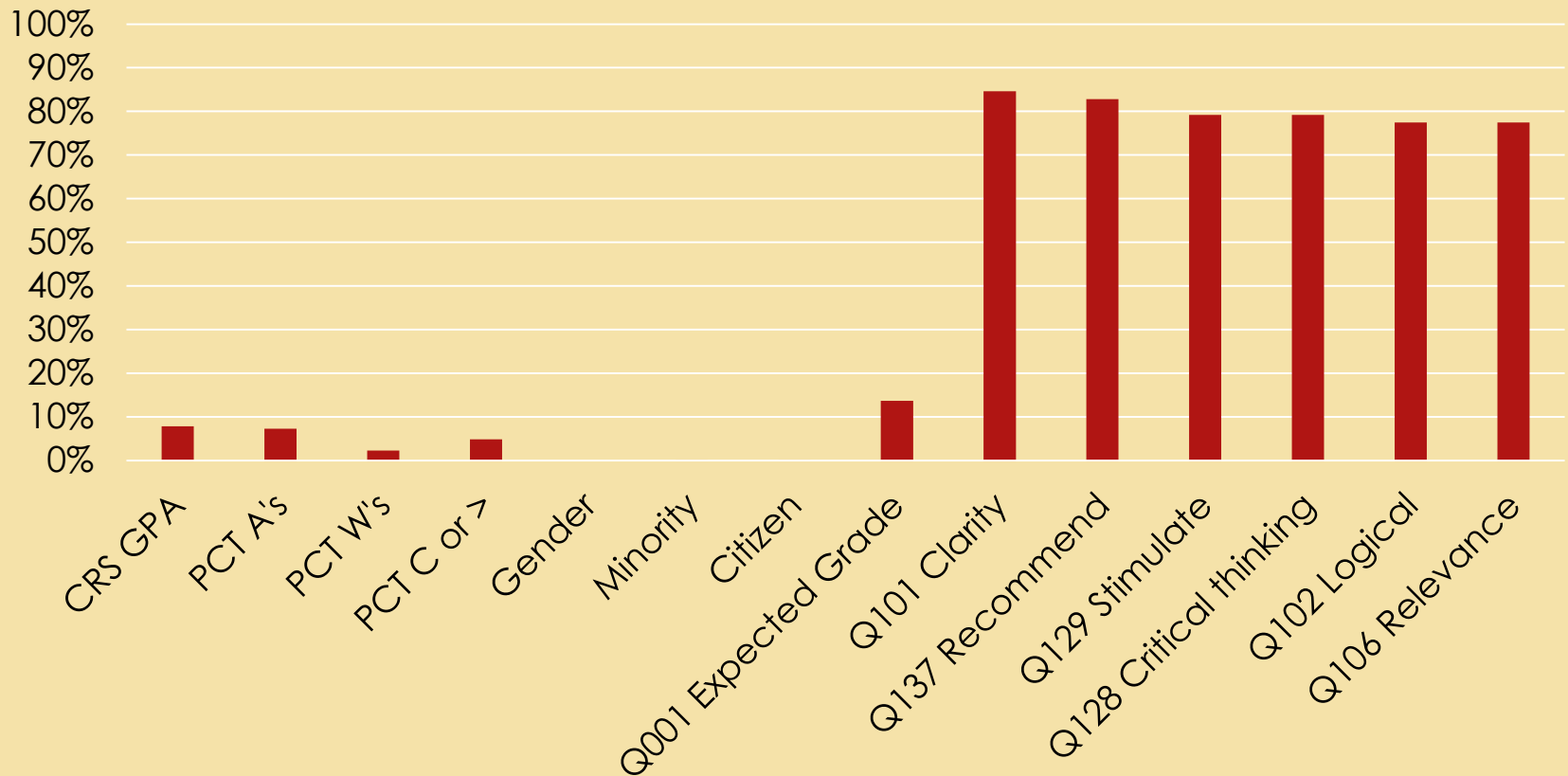
Relationship to other items: Lowest

| Overall, how would you rate the quality of this <u>COURSE</u> ? | Correlation |
|--|-------------|
| 001 What grade do you expect to receive? | 0.37 |
| 003 How would you rate the quality of your effort in this class? | 0.34 |
| 004 What grade were you working to achieve? | 0.27 |
| 002 How often were you fully prepared for class? | 0.24 |
| 005 How often did you attend class? | 0.24 |

The pattern of correlations with other items (what is high versus what is low) supports the measurement of the educational atmosphere of the classroom as the core construct. These items generally relate to elements of course structure.

Example of Shared Variance with Instructor Score

Relative shared variance of elements with Instructor Score
(not unique variance)



Thoughts at this stage ...

- ▶ These data are consistent with others such as Al-Busaidi, Aldhafri and Buyukyavuz (2016), Spooren and Mortelmans (2006), as well as Pietrzak, Duncan and Korcuska (2008). There appears to be a significant single factor that reflects teacher-managed classroom atmosphere and interactions: what Spooren and Mortelmans called *teacher professionalism*.
- ▶ While certainly classes or students can be influenced or “bribed” (people can be in nearly any context), in general these data do not suggest that grades, etc. are a major factor.
- ▶ Thinking of the tool as an **assessment of the educational atmosphere** of the course along a continuum (from assistive to neutral to detrimental to learning) is perhaps best.
- ▶ The research literature further suggests that exposure to feedback alone does not impact performance.
 - ▶ Marsh, H, (2007). Do university teachers become more effective with experience? A multilevel growth model of students' evaluations of teaching over 13 years. *JN of Ed. Psych.*, 99, 775-790)

Questions?

