Data Sources

We surveyed a sample of 177 Georgia Tech instructors from 9 departments. The survey was administered online using ***, and the overall response rate was 66.79% (177/265).

Analyses

- We followed the research methodology of Walter et al.(2016), and compared results of our analysis with their results.

- We first performed exploratory factor analysis (EFA) to examine which our variables can be loaded together. Since the correlation between the factors are expected, we selected a promax rotation, which is the one of the common oblique rotation techniques, to attain an optimal simple structure which attempts to have each variable load on a few factors as possible, but maximizes the number of high loadings on each variable (Gorsuch, 1983). Although Walter et al.(2016) presented the 2F and 5F options for scoring the PIPS, we considered only the 2F option by EFA, to get the simplest model supported by acceptable model fit statistics, and qualitatively logical item groupings.

- (Q : Our data is suitable? I added the multicollinearity part!)

Before running EFA, we examined whether our dataset is suitable for EFA, following An and Sean (2013).

First, we checked if there is a patterned relationship among our variables with the Bartlett's Test of Sphericity. This test shows that we have a patterned relationships if p is lower than 0.0001. We obtained a significant statistics.(p:0.000)

Also, as a follow-up, we checked the absence of multicollinearity with Determinant score. Our determinant score was 0.00002, which is above the rule of thumb of .00001 as an absence of multicollinearity.

Thirdly, we confirmed an acceptable Kaiser-Meyer-Olkin measures (KMO:0.849), which shows that our dataset is suitable for EFA.

- (Q: the requirements for Factor Analysis? Walter et al.(2016) did not show this part!)

To perform a EFA, there has to be multivariate normality within the data (Child, 2006). But, we have left-skewed data. Even though various transformations such as Box-Cox transformation were performed, they were not successful. Instead, we also performed Principal Axis Factor, which is recommended when the data violate the assumption of multivariate normality (Costello & Osborne, 2005).

Also, a factor with EFA should have at least 3 variables. A factor with 2 variables is only considered reliable when the variables are highly correlated with each another($r > 0.7$) but fairly uncorrelated with other variables (Tabachnick & Fidell, 2007). Since we had only 2 factors with more than 3 variables, our analysis was adjustable.

Thirdly, we removed the missing values, and checked the sample size. Comrey and Lee(1992) indicated the ratio of participants to variables should be at least 10:1. So we should have at least 240 because the number of variables is 24. A larger sample size will diminish the error in data and so EFA generally works better with larger sample size (Stevens, 2002).

- (Q: Do PIPS items group together into valid, reliable, and measurable variables?)

Walter et al.(2016) explained three types of validity.

1) Content and face validity: They field-tested the PIPS with a sample of nonparticipating instructors (N=5), and a panel of education researchers at another institution (N=4). But, we only tested the PIPS with Gatech data.
2) Construct Validity (Reliablity): We include respective construct reliabilities in Table 2.

Factor Analysis

We conducted the explanatory factor analysis(EFA). Factors in the 2-factor model include one factor that describes "instructor-centered practice (IC)" (N=9), and another that describes "student-centered practice (SC)" (N=14). In Table 2, unlike Walter et al. (2006), Q6 was removed, and Q17, as well as Q18, were loaded with different factors; Q17 with the factor related with SC practice and Q18 with the factor related with IC practice. Our solution was validated by moderate to good model fit statistics in Table1. We evaluated goodness of fit of our models by using the root-mean-square error of approximation (RMSEA; Steiger, 2000), a comparative fit index (CFI) (Hu and Bentler, 1999), and chi-square/df (Bollen, 1989). The statistics of our model and Walter et al.'s model are comparable.

Table1. PIPS model fit statistics for 2 factors (2F) with Instructor- and Student data

	2F solution
Chi-square(χ^2)	478.31
df	208
Chi-square/ *df*	2.29
CFI	0.737
RMSEA	0.086
N(Number of Observations)	177

Table2. PIPS factor reliability scores, model fit statistics, and items by factor with Instructor- and Student data

	Factor 1:Instructor-centered practice	Factor 2: Student-centered practice
Instructor Data		
Reliability(Instructor/Student data)	0.86	0.72
Number of items	9	14
Eigenvalue	4.002	6.238
Percent variance explained	16.117%	43.498%
Items	Q1, Q3, Q5, Q11, Q18, Q21, Q22, Q23, Q24	Q2, Q4, Q7, Q8, Q9, Q10, Q12, Q13, Q14, Q15, Q16, Q17, Q19, Q20
Maximum possible sum	45	70

- (Q: How to calculate PIPS score?)

This paper used same PIPS score calculation as the one in the Walter et al.2016.

To calculate a factor score from 2F model, we first added scores for the items in that factor, and divide by the maximum possible sum for that factor and then multiply by 100.

Each factor score can vary between 0 (not at all descriptive of my teaching) and 100 (very descriptive of my teaching).

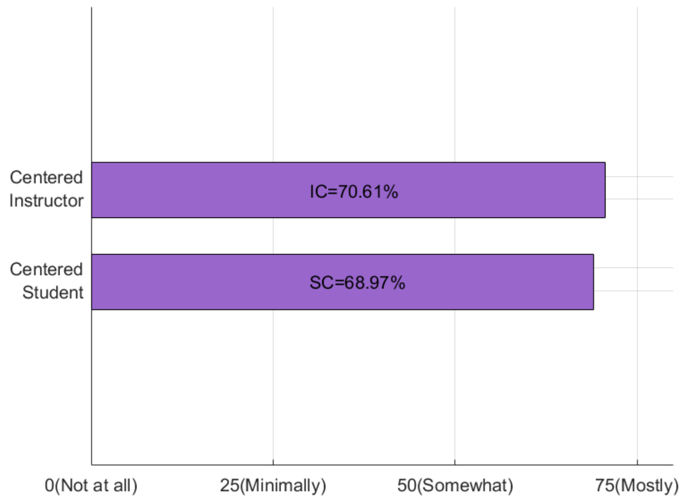
The calculated scores can be utilized to make comparisons among departments, institutions, or demographic groups.

Result

1. PIPS Histogram

PIPS 2 factors are represented on a histogram, which is a frequency-based bar graph. Figure 1 highlights a significant difference between IC- and SC- score. Although we expected that two factors are significantly correlated, no significant correlation between IC- and SC- score(r= -0.0714, p= 0.3253) supports the independent nature of these factors, which is consistent with the finding in Walter et al. (2016). Walter et al. (2016) suggested that this is because the 2F item loadings for a varimax rotation, which is used when the factors are rotated 90 from each other, are equivalent to the ones on a promax rotation.

FIGURE 1. Averaged PIPS 2F Scores for Instructors



2. PIPS Scatter Plot

Figure 2 presents the scatter plot of individual 2F PIPS scores for instructors in various sampled departments. There were total of 192 instructors (N=192) dividing into multiple departments (N: AE=20, BME=45, CHBE=10, CEE=26, ECE=45, ISYE=13, MSE=4, ME=24, OTHERS=5). 15 data sets were duplicated for the instructors who were in two departments. Figure 3 shows the mean value of the departments in instructor-centered scores from Figure 2. Significant differences were found in the mean score of each department for the SC practice (p: 0.0106). However, the significant difference did not appear in the instructor's IC practice (p: 0.5641). As shown in Figure 3, BME is depicted as the highest SC practice score. In contrast, CHBE appears to be the highest IC practice score.

(There is no information about which department has the highest- and the lowest SC,IC scores in Walter.et. al(2016). So we are not able to compare our results with the ones in Walter. et. al(2016))

Figure2. PIPS Scores for Instructors in the 9 Sampled Departments

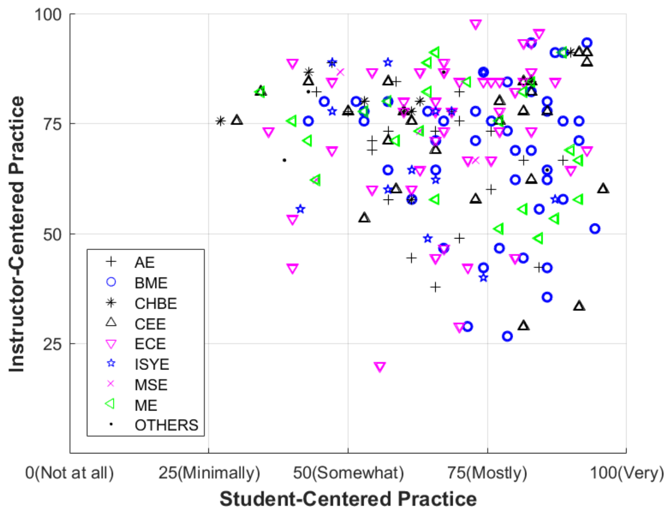
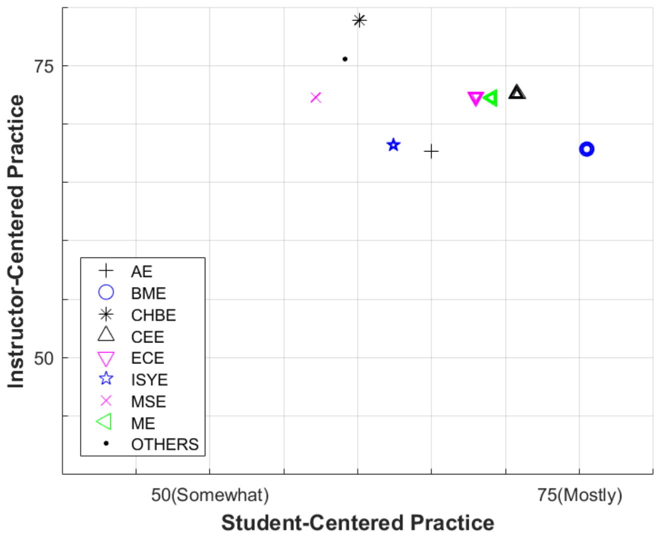


Figure3. The Averaged PIPS Scores for Instructors in the 9 Sampled Departments



In figure 4, the points on the scatter plot represent PIPS 2F scores are based on courses in each department. There were total of 182 instructors (N=182), according to courses in the same department (N: AE=17, BME=40, CE=3, CEE=16, CHB=10, COE=13, ECE=43, ISYE=12, MSE=3, ME=22, MP=1, NRE=2). If the number of the course is below 3, it was eliminated. There were duplicate data for the instructor who was in two departments (N=5). The figure 5 shows the mean value of instructor- and student-centered scores in the respective departments from the figure 4. Significant differences were found in the mean score of SC- and IC- practice score (SC practice (p: 0.0023), IC practice (p: 0.0347)). In Figure 5, while BME is depicted as the highest SC practice score, CHBE appears to be the highest IC practice.

(Analysis for courses was newly added. So we are not able to compare our results with the ones in Walter. et. al (2016))

Figure4. PIPS Scores for Courses in the Each Department

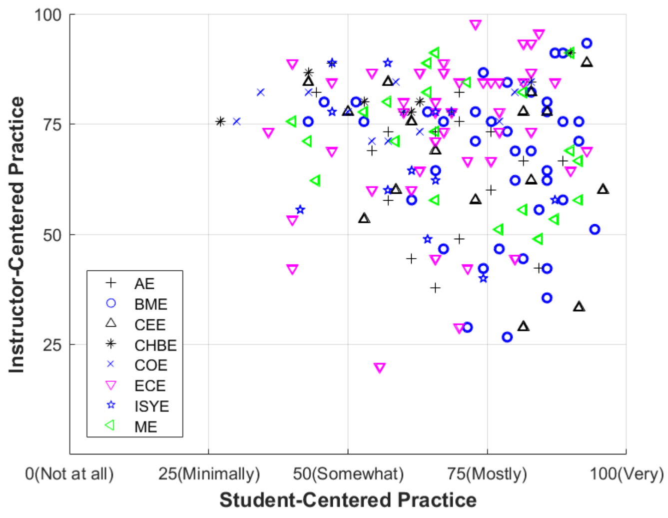
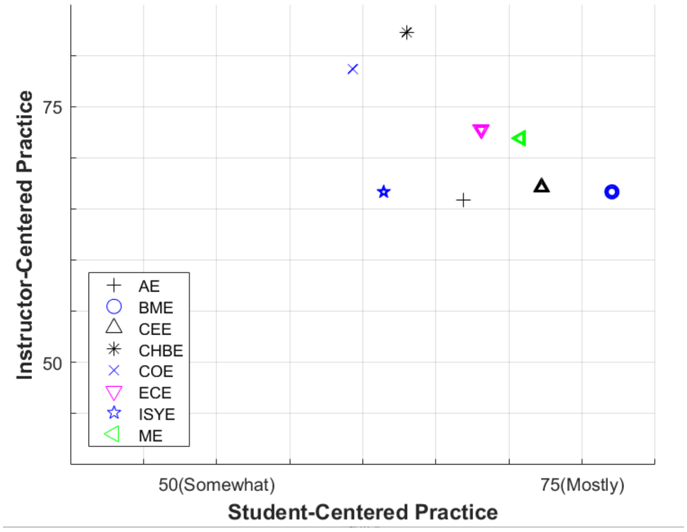


Figure5. The Averaged PIPS Scores for Courses in the Each Department



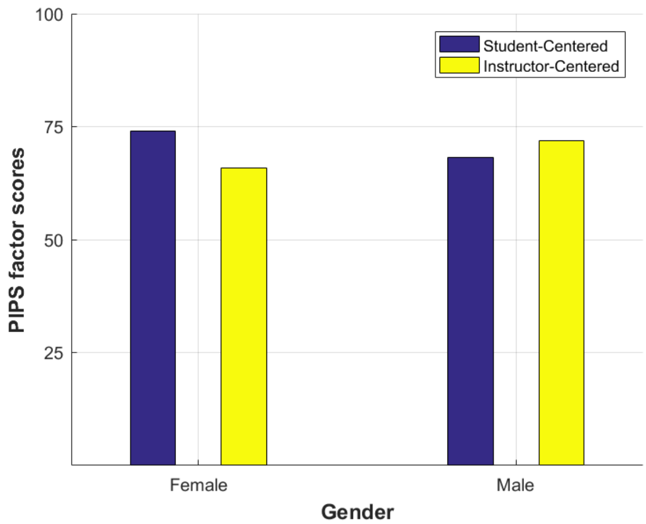
3. Demographic Differences

We examined whether there exists demographic differences in SC and IC scores, as generated by the 2F PIPS model. We ran two sample t-tests and ANOVA comparisons to explore demographic differences between and among PIPS scores for different instructor groups (Gender, Tenure, Rank, Years teaching, Years in the higher education).

1) By Gender

We explored the instructor data (N=26 female; N=139 male; N=12 no answer) to verify whether there exists a significant difference. Walter et al. (2016) found significant differences in 2F PIPS scores between genders. In contrast, there was no significant difference between gender for the mean of the IC (p: 0.4711), SC (p: 0.7124) in our data, even though they seemingly have a difference.

Figure6. The Averaged PIPS scores by Gender



2) How time is spent in class: doing lecture, small group, individualized instruction, and other instruction

We found that the lecture-based pedagogies were negatively correlated with SC practice, and positively with IC practice. For the small group, instructors felt that there was a negative correlation with IC, but a positive correlation with SC. These results are consistent with the findings in Walter et al. (2016), except the individual instruction part.

Table2. Pearson Correlations among PIPS, and Participant Estimations of How time is spent in Class: Doing Lecture, Small Group Work, Individualized Instruction, and Other Instruction (Instructor)

	IC	SC	Lecture	Small group	Individual	Others
SC	-0.0847 (0.2621)	1	-0.5357** (1.5584e-14)	0.5855** (1.1440e-17)	0.0238 (0.7530)	0.0338 (0.6556)
IC	1		0.4689** (4.6014e-11)	-0.4102** (1.4323e-08)	0.0602 (0.4262)	-0.3040** (3.8921e-05)

The list of 'Others': (Number is the percentage of the other based pedagogies)

100	Combination of large and small group meetings/work
90	Instructor programs at front of class while students 'code along' on their laptops. Instructor prompts students for 'next step' in code frequently, and fields questions/suggestions
70	meeting of instructor with student design teams
34	Problem session
34	Presenting Material
34	Active problem solving in groups with instructor support
33	the students did in-class labs once per week
30	problems sets and small projects
30	Group Projects
30	Instructor facilitated class discussions
25	Answering questions asked by me
25	Presenting
25	class discussion
20	term project reports
20	working problems as a class
20	class discussion
20	Discussion
20	Team Meeting
16.67	Class discussion
15	Discussing problems posed by instructor
15	Discussion
15	office hours
15	misc. subteam activities
15	whoel class discussions
12	Weekly quizzes
10	interactive questions/discussions with students
10	Activities in the computer lab
10	Client visits
10	peer assessments
10	technical demos, tours
10	discussions with me
10	Consult with other faculty
10	Quizzes
10	journal article review
10	Taking tests and quizzes
10	discussion of topics
10	quizzes and practice quizzes
10	quiz
10	extra problem-solving, self-introspection
10	Student questions/discussions
8	Class presentations
5	on some outside work they do
5	presenting material
3	asking questions
2.5	taking quizzes
2.5	Students interacting with demos

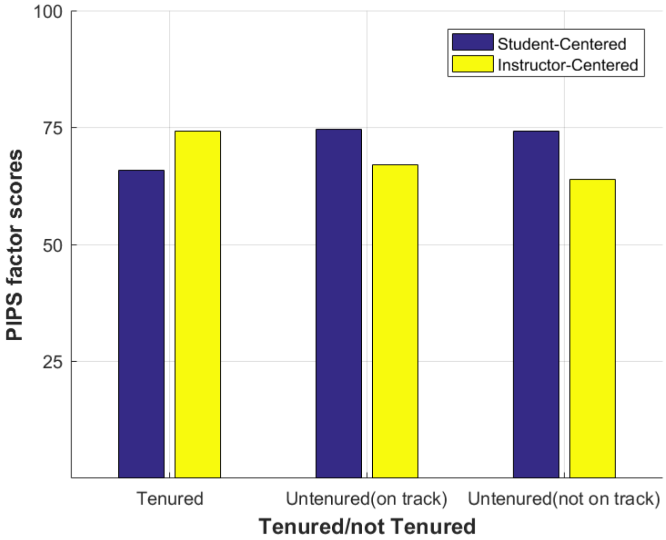
3) By Tenure or not.

Data was divided into three groups (Tenured: 109, Untenured (on the Tenured track): 28, Untenured (not on the Tenured track): 40) There were significant differences of mean scores between three divisions for SC and IC practice. For the student-centered practice, Tenured instructors had significantly lower score of the averaged PIPS than untenured (on the track) (p: 0.0189), and untenured instructors (p: 0.0079). But, the averaged PIPS score of untenured instructors (on the track) was not able to be compared with the one of the untenured (not on the track) (p: 0.9952).

For the instructor-centered practice, Tenured instructors had significantly higher scores than untenured instructors (not on the track) in IC practice (p: 0.0012), but the others were not able to compared to each other (the untenured instructors (on the tenure track) with the untenured (not on the tenure track) (p: 0.7019), the tenured with the untenured (on the tenure track) (p: 0.0836))

(Analysis for courses was newly added. So we are not able to compare our results with the ones in Walter. et. al (2016))

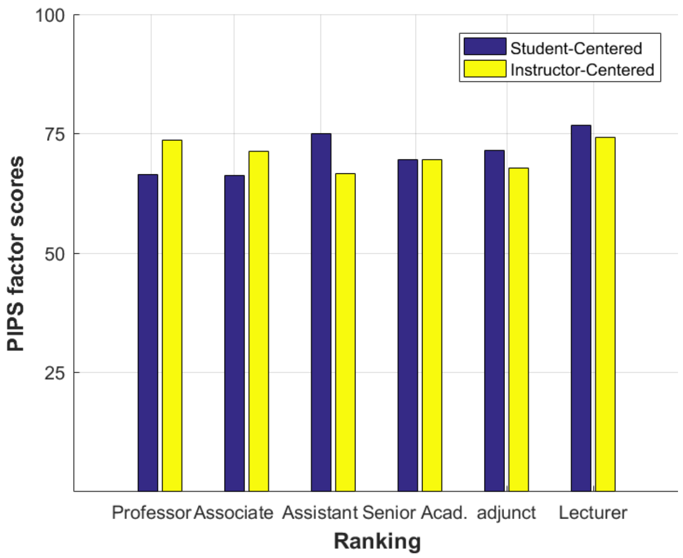
Figure7. The Averaged PIPS Scores by Tenure or Not



4) By Ranking

The data is following: (N=76 Professor; N=28 Associate Professor; N=26 Assistant Professor; N=9 Senior Academic Professor; N=11 Adjunct or Par-time; N=7 Lecturer) There was no significance of mean scores by ranking for SC (p: 0.1024) and IC practice (p: 0.2025). By contrast, significant differences in 2F PIPS scores were founded among faculty of differing academic rank in Walter et al. (2016)

Figure8. The Averaged PIPS Scores by Ranking



5) By Years Teaching

We also conducted the correlation analyses to examine the 2F PIPS factors relationships to class size (N=177), years teaching (N=177), and years at the institution (N=177). Table4 represents that years teaching were negatively correlated with SC factor. This result is consistent with Haiva (2000) and Kuh et al. (2004), which suggested that more senior faculty are often less innovative than younger faculty. However, years teaching were not correlated with IC factor. These results contradicted the findings of Water et al. (2016) where they showed the years of teachings are correlated with IC factor, not with SC factor.

6) By Years in the Higher Education

Similar to the above results, while years in the higher education were negatively correlated with SC factor, they were not correlated with IC factor. These results also contradicted the findings of Water et al. (2016) where they showed the years in the higher education are correlated with IC factor, not with SC factor.

Table4. PIPS Factor Correlations with Class Size, Years Teaching, and Years in the Higher Education

	Years teaching(r(p))	Years in higher education (r(p))
Student-centered	-0.3398**(4.7739e-06)	-0.3231**(1.4479e-05)
Instructor-centered	0.0905(0.2363)	0.1157(0.1296)
Years teaching	1	0.8741** (1.6533e-55)
Years in higher education		1

(We do not have a survey on class size. So, it is not able to be compared.)

Implications for Policy

in Walter et al. (2016)

…It is important for researchers, institutions, and policy makers to have a valid and reliable instrument that can describe a range of traditional and research-based teaching practices across instructors from multiple departments. This can be useful, for example, to identify outlier departments (positive deviants that can be learned from) or to document the results of change initiatives longitudinally….