

Characterizing Professors and Courses Based on Student Perceptions

*L. T. Upchurch, (leotup@aol.com), *F. Wu (wuf@mytu.tuskegee.edu)
and **T.L. Upchurch-Poole (tlpoole@crimson.ua.edu)

*College of Business and Information Science, *Tuskegee University, Tuskegee,
Alabama 36088

** College of Education, University of Alabama, Tuscaloosa, Alabama 35487

ABSTRACT

This paper characterizes professors and courses based on student perceptions. Whatever student evaluators of professors and courses perceive is most certainly latent (feelings and ideas not easily articulated). How a student feels about a given professor/course experience is not adequately captured in a binary sense (“The professor was my cup of tea” /the professor was the pits” similarly for the course). We create and employ a device known as a domain defining set (DDS) which offers a mechanism for addressing the latency problem. Primary objectives include examination of differences in perceptions across disciplines, e.g., natural sciences, social sciences, engineering, and computer science and especially the development of a measure herein labeled ‘**IKOFF**’ (interesting, knowledgeable, organized, fair, and friendly). This measure is a composite of the adjectives, attributes, qualities, traits, characteristics [**AADQTCs**] that student evaluators most often use to capture their feelings about courses and professors / teachers. Additionally, an instrument linked to the five-factor theory of personality assessment (**FFTP**) is employed to try to gauge tendencies of student evaluators. Use of ‘online’ methodology has facilitated this work, particularly in ease of gathering data.

KEY WORDS

Domain Defining Set (DDS), IKOFF, Student Evaluator, Five Factors Theory of Personality (FFTP).

INTRODUCTION AND BACKGROUND

The academic climate in colleges and universities today requires students to assess the experiences they have had in various courses and with corresponding professors /teachers. In this work, we first create a device that allows students (evaluators) to express their perceptions of the extent (amount)to which a set of attributes , adjectives, descriptors, , qualities, traits, characteristics (**AADQTC**), are observable with respect to the course and with respect to the professor/teacher. We denote the set (**AADQTC**) as a domain defining set (**DDS**). Hence we have set about characterizing courses and professors/teachers.

It is useful to specify exactly how this device which we have labeled ‘DDS’ is constructed. An example is the clearest way forward. Suppose that at BasketWeaving College, every freshperson is required to take SANSKRIT 109. Professor Ravendiba, the only SANSKRIT professor, has gotten his wish of 3-sections, 3-days per week. The 147 first- year students are divided nearly evenly over the 3-sections (47, 49, 51).

Never before has there been such buzz about Ravendiba, so much so that the Dean of BasketWeaving College has been admonished by the President to, “investigate the Ravendiba matter”. In consultation with Ravendiba, it is decided that 15-minutes of the next meeting of each section of SANSKRIT 109 can be used to ‘elicit a domain defining set (DDS). With Ravendiba out of the room, each student is asked to ‘text’ to a number

supplied by the Dean, 5-one word or hyphenated word (AADQTC) that seem most nearly to capture their impression (perception) of Ravendiba . Likewise, text 5-one word or hyphenated word (AADQTC) that seem most nearly to capture their impression (perception) of the course.

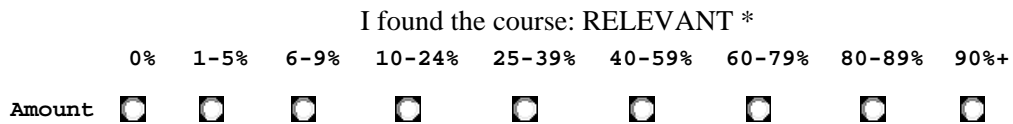
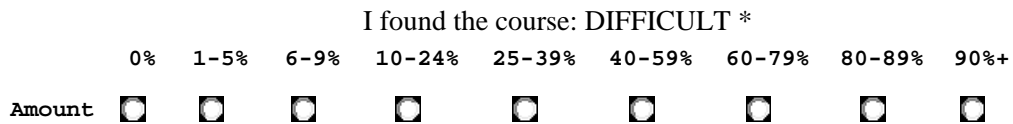
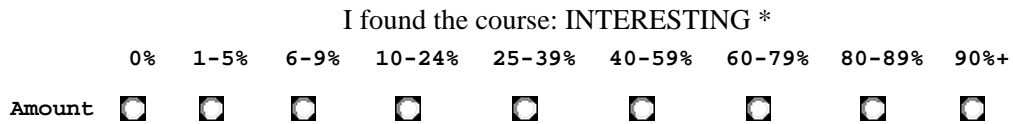
When the Dean returns to his office later that day, he finds that his clerical assistant has ‘assembled two joint domain defining sets’ one for the **COURSE**, one for the **PROFESSOR (DDS_c, DDS_p)**. **DDS_c** could have $[5(47) + 5(49) + 5(51)] = 735$ words; **DDS_p** could have $[5(47) + 5(49) + 5(51)] = 735$ words . This activity has grossly outlined the sum of the experiences of all 147 students.

The next phase calls for refining each DDS (**DDS_c, DDS_p**). Remove (AADQTC) that are redundant. It should not be surprising to find student assessors giving the same (AADQTC). Theoretically the (AADQTC)-subset remaining could be large, but in practice, 30 to 50 items should be more than adequate to cover the most useful and revealing (AADQTC).

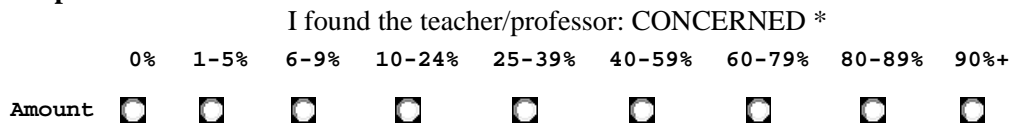
Now the ‘Characterization Process’ of course and professor/teacher can begin. Here is a brief look at the device. It is administered online and the data, student responses to the 33 course / professor related items and 10 other items that attempt to assess (infer) how some physical behaviors (tendencies), thought to link with personality, might be related to the 33 primary items. Below, we provide subsets of both the COURSE (DDS_c) and PROFESSOR (DDS_p) DDS.

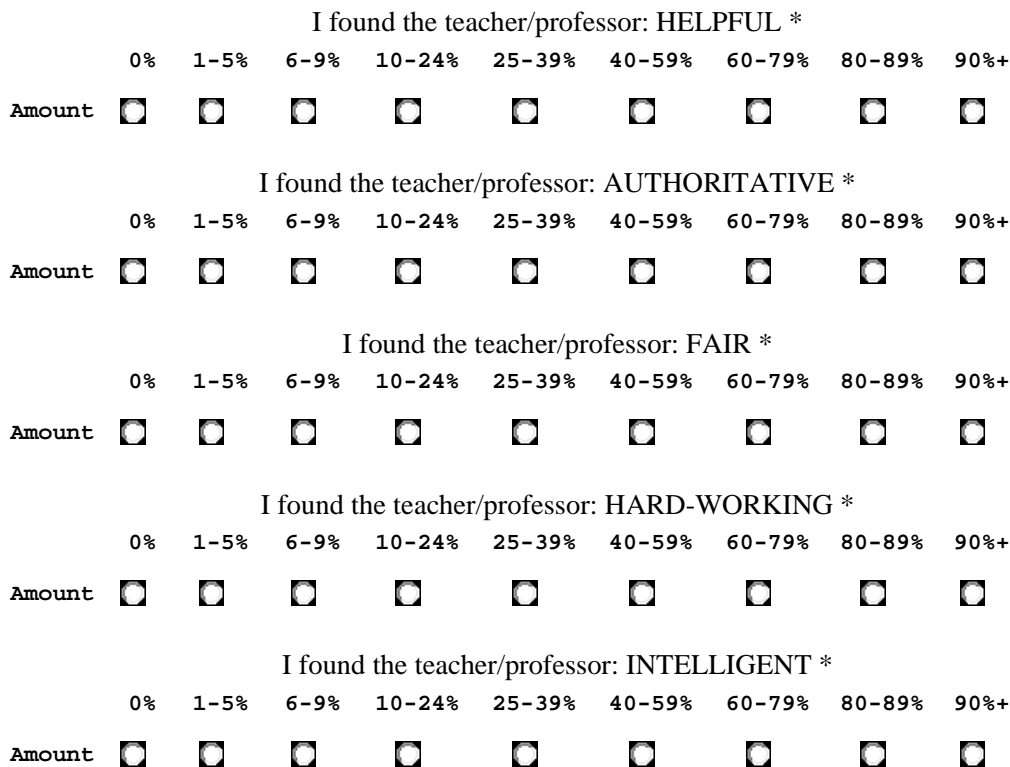
Five COURSE and **28 PROFESSOR DDS** items were presented to respondents (students) [**COURSE:** interesting, informative, applicable, difficult, relevant, and ‘respondent-contributed’; **PROFESSOR:** concerned, helpful, authoritative, fair, hard-working, intelligent, consistent, conscientious, detailed, thorough, receptive, prepared, interesting, experienced, perceptive, friendly, and ‘respondent-contributed’] in the following format:

DDSc



DDSp





ASSESSING THE ASSESSOR

We have used a, **ten (10) item INSTRUMENT** designed to gauge respondent personality tendencies in conjunction with the DDS_j items. **[These PHYSICAL BEHAVIORS are thought to link with ‘personality’]**. For clarity, we provide a few of the items included in the INSTRUMENT.

You usually walk *

- fairly fast, with long steps
- fairly fast, with little steps
- less fast, head up, looking the world in the face
- less fast, head down
- very slowly

When talking to people, you *

- stand with your arms folded
- have your hands clasped
- have one or both your hands on your hips
- touch or push the person to whom you are talking
- play with your ear, touch your chin, or smooth your hair

When something really amuses you, you react with *

- a big, appreciative laugh
- a laugh, but not a loud one
- a quiet chuckle
- a sheepish smile

These ten items relate to the ‘five-factor theory of personality’, acronym: **OCEAN** (openness, conscientiousness, extraversion, agreeableness, neuroticism). It is claimed by some that these items broadly capture one's 'approach to the world'. Each of the ten (10) items, respectively have associated sub-categories, each sub-category having assigned to it a numerical score ranging between 0 and 7. The summated subcategories over the ten items produces an **OVERALL personality test score (PT_SCR)**. Possible **PT_SCRs** ranged from high teens to a maximum of 60.

EXPERIMENTAL RESULTS

The procedure outlined in the contrived example at the outset of this exposition was used to elicit the DDS which produced the data shown below in summarized form (descriptive statistics). A total of 103 students in 47 different courses, representing more than 25 different majors responded to the survey. We have chosen to categorize the courses as STEM (47) and NON-STEM (56). The sample used in this study was comprised of 60 females and 43 males. Academically, respondents were classified as freshman (14), sophomores (19), juniors (17), seniors (19), masters (7), and doctoral (26). Participants in this study came from more than 40 different universities, in more than 30 states and Canada. For the 33 primary items (5 –course AADQTC_j and 28- professor AADQTC_j), survey respondents chose responses from 9 -intervals [0, 1-5, 6-9, 10-24, 25-39, 40-59, 60-79, 80-89, 90+]. Intervals chosen by respondents expressed their perceptions of the amount of an AADQTC_j.

Observe that the first 10-rows show data for AADQTCs associated with the course (DDS_i). Rows-1 & 2 contains the respective assessed means for the amount of the perceived quality ‘INTERESTING’ (73.83) and the mean of the transformed value ‘INTERESTING_TFN’ (1.1048). Each of the 103 observations for ‘INTERESTING’ was transformed by application of the square root transformation:

$$Y_i = [(Y_i/100) + .5]**.5$$

The same procedure was applied to every AADQTC. Application of this particular transformation was appropriate because possible responses ranged from 0 – 100%.

An interesting question is whether differences exist among some AADQTCs (particularly the ‘IKOFF’ AADQTCs’) when certain demographics (e.g. gender, subject matter category [STEM, NON-STEM], and class [doctorate, masters, senior, junior, sophomore, freshmen]), are simultaneously considered. This consideration yields an *ex post facto* two-factor experimental design (e.g., a 2 X 7 in the case of gender X AADQTC_j; a 2 X 7 in the case of subject matter X AADQTC_j; or a 6 X 7 in the case of class X AADQTC_j). There is also the question of first-order interaction. The two considerations (ex post facto 2-factor experimental design and the likelihood of factor level interaction) suggest that the general linear model (GLM) is very appropriate.

Now, we turn to our first priority: characterization of professors and courses. A summary characterization for each professor is given by the **IKOFF** score. **‘IKOFF’ SCORE DERIVATION:** We compute an IKOFF score for each professor

(teacher) by summing the **DDS**-items comprising '**IKOFF**': [INTERESTING, KNOWLEDGEABLE, ORGANIZED, FAIR, and FRIENDLY].

Descriptive Statistics FOR DDS (DDS_c, DDS_p)

DDS_c [SUBSET]

Variable	Mean	SE Mean	StDev	Min.	Q1	Median	Q3	Max.
INTERESTING	73.83	2.63	26.73	3.00	69.50	84.50	95.00	95.00
INTERESTING_TFN	1.1048	0.0131	0.1334	0.728	1.0932	1.1597	1.2042	1.2042
INFORMATIVE	79.00	2.55	25.91	3.00	84.50	95.00	95.00	95.00
INFORMATIVE_TFN	1.1288	0.0126	0.1283	0.728	1.1597	1.2042	1.2042	1.2042
APPLICABLE	76.78	2.66	26.99	3.00	69.50	84.50	95.00	95.00
APPLICABLE_TFN	1.118	0.0133	0.1345	0.728	1.0932	1.1597	1.2042	1.2042
DIFFICULT	52.49	3.26	33.10	0.00	17.00	49.50	84.50	95.00
DIFFICULT_TFN	0.9982	0.0167	0.1697	0.7071	0.8185	0.9975	1.1597	1.2042
RELEVANT	76.24	2.56	25.99	0.00	69.50	84.50	95.00	95.00
RELEVANT_TFN	1.1164	0.0125	0.1272	0.7071	1.0932	1.1597	1.2042	1.2042

The above table shows summary descriptive statistics for a subset of **5-DDS_c** course items (in percent, followed immediately by the transformed value). Similarly, the table below accomplishes the same thing for a subset of **10- DDS_p** professor items.

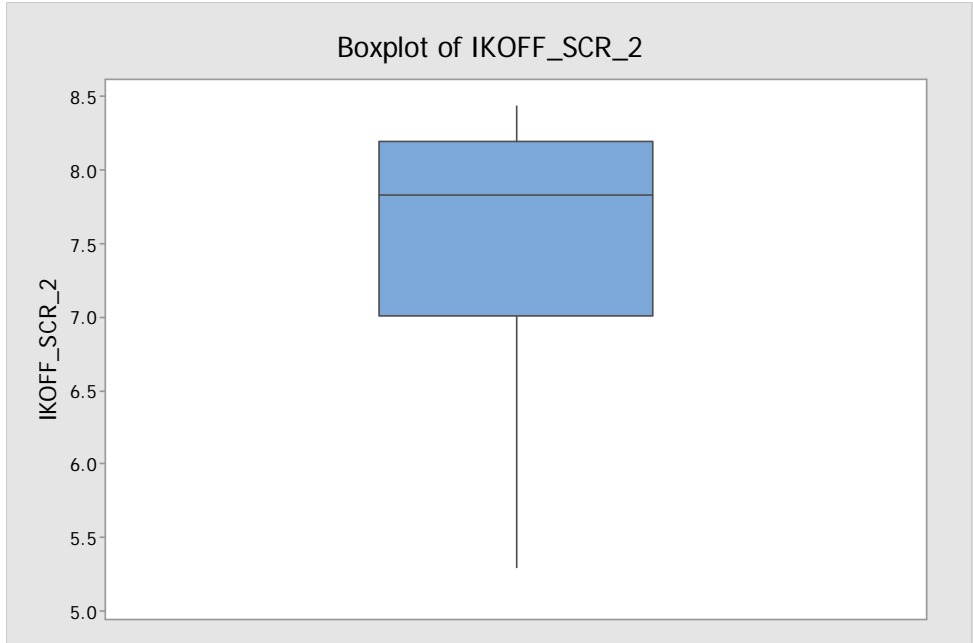
DDS_p [SUBSET]

Variable	Mean	SE Mean	StDev	Min.	Q1	Median	Q3	Max.
CONCERNED	75.56	2.90	29.39	0.00	69.50	84.50	95.00	95.00
CONCERNED_TFN	1.1107	0.0146	0.1484	0.7071	1.0932	1.1597	1.2042	1.2042
INTERESTING_P	77.57	2.49	25.23	3.00	69.50	84.50	95.00	95.00
INTERESTING_P_TFN	1.1227	0.0122	0.1242	0.728	1.0932	1.1597	1.2042	1.2042
FRIENDLY	80.84	2.49	25.23	0.00	84.50	95.00	95.00	95.00
FRIENDLY_TFN	1.1371	0.0123	0.1247	0.7071	1.1597	1.2042	1.2042	1.2042
INTERESTING_1	73.83	2.63	26.73	3.00	69.50	84.50	95.00	95.00
INTERESTING_TFN_1	1.1048	0.0131	0.1334	0.728	1.0932	1.1597	1.2042	1.2042
INFORMATIVE_1	79.06	2.55	25.91	3.00	84.50	95.00	95.00	95.00
INFORMATIVE_TFN_1	1.1288	0.0126	0.1283	0.728	1.1597	1.2042	1.2042	1.2042
APPLICABLE_1	76.78	2.66	26.99	3.00	69.50	84.50	95.00	95.00
APPLICABLE_TFN_1	1.118	0.0133	0.1345	0.728	1.0932	1.1597	1.2042	1.2042
DIFFICULT_1	52.49	3.26	33.10	0.00	17.00	49.50	84.50	95.00
DIFFICULT_TFN_1	0.9982	0.0167	0.1697	0.7071	0.8185	0.9975	1.1597	1.2042
RELEVANT_1	76.24	2.56	25.99	0.00	69.50	84.50	95.00	95.00
RELEVANT_TFN_1	1.1164	0.0125	0.1272	0.7071	1.0932	1.1597	1.2042	1.2042
CONCERNED_1	75.56	2.90	29.39	0.00	69.50	84.50	95.00	95.00
CONCERNED_TFN_1	1.1107	0.0146	0.1484	0.7071	1.0932	1.1597	1.2042	1.2042
HELPFUL_1	79.28	2.48	25.20	7.50	84.50	95.00	95.00	95.00
HELPFUL_TFN_1	1.1303	0.0122	0.1239	0.7583	1.1597	1.2042	1.2042	1.2042
PT_SCR_1	42.417	0.501	5.085	30.00	38.00	43.00	47.00	54.00

Because respondents expressed their perceived amounts of the AADQTC_j in percent intervals, we have chosen to use mid-points of the '8 amount' intervals following a possible '0 amount' for a given AADQTC_j [0, 1-5%,6-9%, 10-24%, 25-39%, 40-59%, 60-79% 80-89%, 90-100%].

For purposes of this present effort, we concentrate on ‘professor characterization’ using **IKOFF scores** only. Graphically, these scores are presented below.

IKOFF SCORES FOR 103 PROFESSORS



Descriptive Statistics: IKOFF_SCORE

Variable	Min.	Q1	Median	Q3	Max.	Skewness	Kurtosis
IKOFF_SCR_2	5.287	7.00	7.832	8.19	8.429	-1.19	0.47

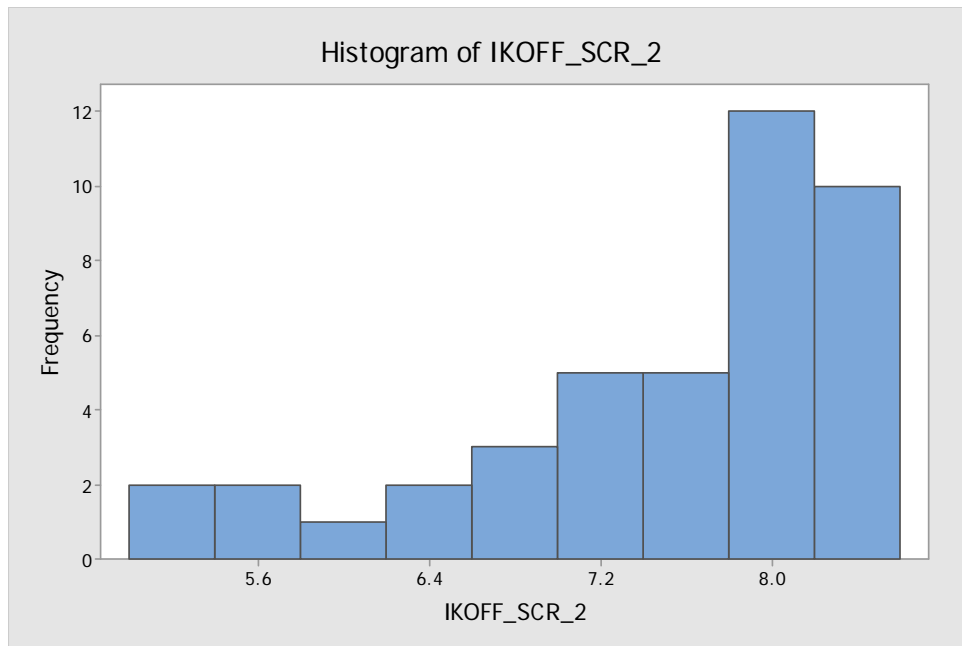
Descriptive Statistics: INTERESTING [COURSE], GENDER

Results for STEM = NO

Variable	GENDER	Mean	SE Mean	StDev	Min.	Q1	Median	Q3	Max.
INTERESTING	Female	75.84	4.16	24.63	7.50	69.50	84.50	95.00	95.00
	Male	75.32	6.03	28.28	3.00	69.50	84.50	95.00	95.00

Results for STEM = YES

Variable	GENDER	Mean	SE Mean	StDev	Min.	Q1	Median	Q3	Max.
INTERESTING	Female	73.15	5.12	26.60	3.00	69.50	84.50	95.00	95.00
	Male	69.37	6.94	30.26	3.00	49.50	84.50	95.00	95.00



SELECTED IKOFF VARIABLES IN RELATION TO 'STEM' INTERESTING [COURSE] & FRIENDLY [PROFESSOR]

INTERESTING	STEM		TOT
	No	Yes	
LOW	7	6	13
MED	13	14	27
HIGH	37	26	63
TOT	57	46	103

Chi-Square Test for Association:

INTERESTING	STEM		TOT
	NO	YES	
LOW	7	6	13
	7.19	5.81	
MEDIUM	13	14	27
	14.94	12.06	
HIGH	37	26	63
	34.86	28.14	
TOT	57	46	103

Pearson Chi-Square = 0.870, DF = 2, P-Value = 0.647

Likelihood Ratio Chi-Square = 0.868, DF = 2, P-Value = 0.648

**Chi-Square Test for Association:
IKOFF VARIABLE: FRIENDLY**

PROFESSOR FRIENDLY	STEM		TOT
	NO	YES	
LOW	8	3	11
	6.087	4.913	
MEDIUM	7	6	13
	7.194	5.806	
HIGH	42	37	79
	43.718	35.282	
TOT	57	46	103

Pearson Chi-Square = 1.509, DF = 2, P-Value = 0.470

Likelihood Ratio Chi-Square = 1.575, DF = 2, P-Value = 0.455

There is a prevailing perception that 'STEM CLASSES' are more challenging and that professors in these classes are 'more no-nonsense' (translated less 'sociable / less friendly').

While we do not find statistically significant association ($\alpha = .05$) between perceived level of course difficulty and whether or not a course is or is NOT STEM, the level of significance achieved ($p = 0.11$) is MOST DEFINITELY 'heading in the direction' of our hypothesis (e.g. STEM classes will possess a greater tendency to be characterized as DIFFICULT ($\chi^2 = 4.41$ with 2-degrees of freedom; and corresponding $p = 0.11$)).

**Chi-Square Test for Association:
Course quality: Difficult vs STEM_NOT_STEM**

COURSE DIFFICULT _LVL	STEM		TOT
	NO	YES	
LOW	29	14	43
	23.80	19.20	
MEDIUM	12	13	25
	13.83	11.17	
HIGH	16	19	35
	19.37	15.63	
TOT	57	46	103

Pearson Chi-Square = 4.405, DF = 2, P-Value = 0.111

Likelihood Ratio Chi-Square = 4.465, DF = 2, P-Value = 0.107

Table: INTERESTING [COURSE], ACAD_CLASS

Rows: INTERESTING (AMOUNT)	Columns: ACAD_CLASS						
	Doctorate	Freshman	Junior	Master	Senior	Sophomore	All
3.0	0	2	1	0	1	0	4
7.5	0	0	0	0	1	0	1
17.0	0	1	1	1	2	1	6
32.0	1	0	1	0	0	0	2
49.5	2	1	3	0	1	2	9
69.5	8	2	2	0	3	3	18
84.5	5	3	2	3	7	5	25
95.0	12	4	7	3	4	8	38
All	28	13	17	7	19	19	103

Table: DIFFICULT, ACAD_CLASS [COURSE]

DIFFICULT Rows (AMOUNT)	Columns: ACAD_CLASS						
	Doctorate	Freshman	Junior	Master	Senior	Sophomore	All
0.0	1	1	0	0	1	0	3
3.0	2	1	0	0	1	2	6
7.5	1	0	2	0	2	1	6
17.0	2	3	1	2	2	5	15
32.0	4	2	3	0	3	1	13
49.5	1	2	2	3	1	0	9
69.5	4	1	1	1	2	7	16
84.5	10	1	4	0	3	2	20
95.0	3	2	4	1	4	1	15
All	28	13	17	7	19	19	103

Table : INTERESTING_P, ACAD_CLASS [PROFESSOR]

INTERESTING_P Rows(AMOUNT)	Columns: ACAD_CLASS						
	Doctorate	Freshman	Junior	Master	Senior	Sophomore	All
3.0	0	0	0	1	0	0	1
7.5	0	0	0	1	1	0	2
17.0	1	0	1	0	3	1	6
32.0	0	1	1	1	0	0	3
49.5	3	1	0	1	1	0	6
69.5	1	2	3	0	2	6	14
84.5	5	3	1	2	6	6	23
95.0	18	6	11	1	6	6	48
All	28	13	17	7	19	19	103

Table : FRIENDLY, ACAD_CLASS [PROFESSOR]

FRIENDLY
Rows(AMOUNT)
Columns: ACAD_CLASS

	Doctorate	Freshman	Junior	Master	Senior	Sophomore	All
0.0	0	0	0	1	0	0	1
7.5	0	0	0	1	2	0	3
17.0	1	0	2	0	1	0	4
32.0	0	2	0	0	1	0	3
49.5	1	1	1	0	1	1	5
69.5	0	0	3	1	2	2	8
84.5	5	1	4	1	3	2	16
95.0	21	9	7	3	9	14	63
All	28	13	17	7	19	19	103

Cell Contents: Count

Below, we create a bi-variate table by collapsing the row variable FRIENDLY_(AMOUNT) into 3 categories and the column variable ACADEMIC CLASS into 4 categories. Performing the Chi Square Test of association on the resulting 3 X 4 table gives a Chi Square value of 5.70 and a Likelihood Ratio Chi-Square of 6.12, each with DF = 4 and corresponding *p-values*, respectively of 0.220 and 0.190. Once again, we do not achieve statistical significance at the $\alpha = .05$ level. Nevertheless, the values achieved are encouraging.

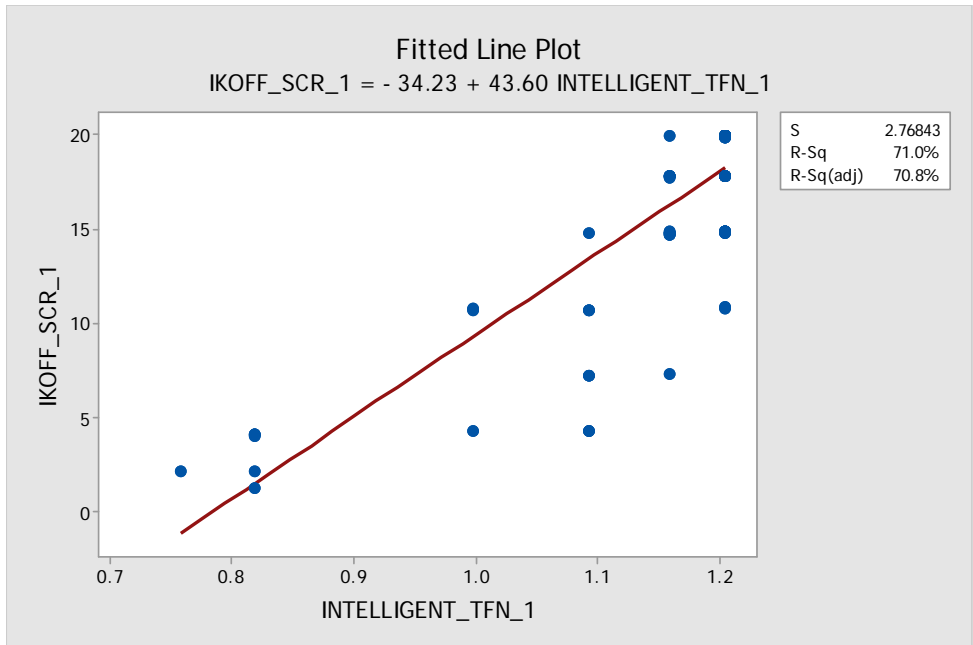
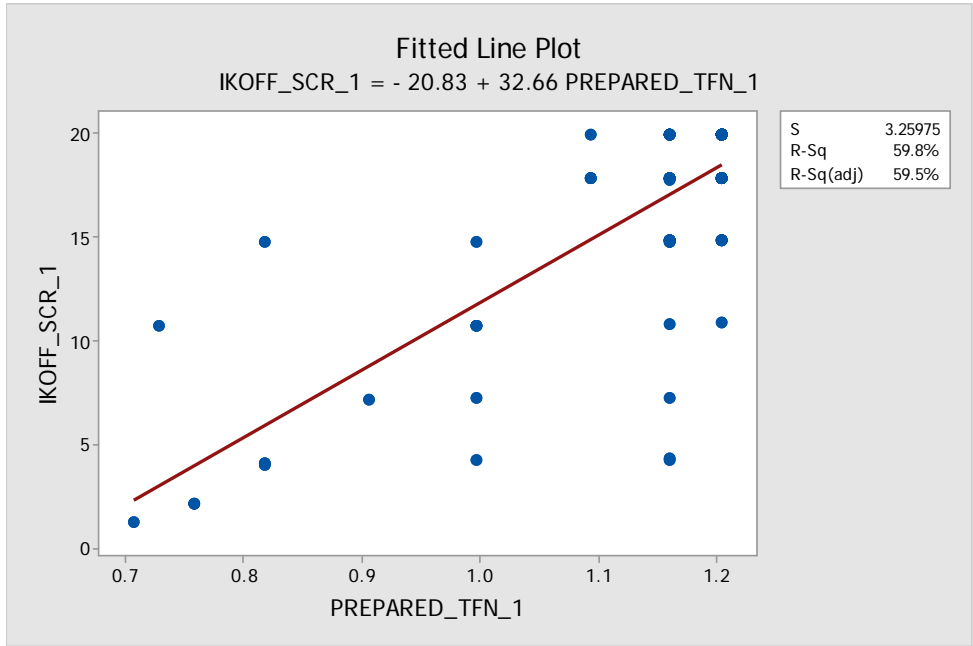
Chi-Square Test for Association:

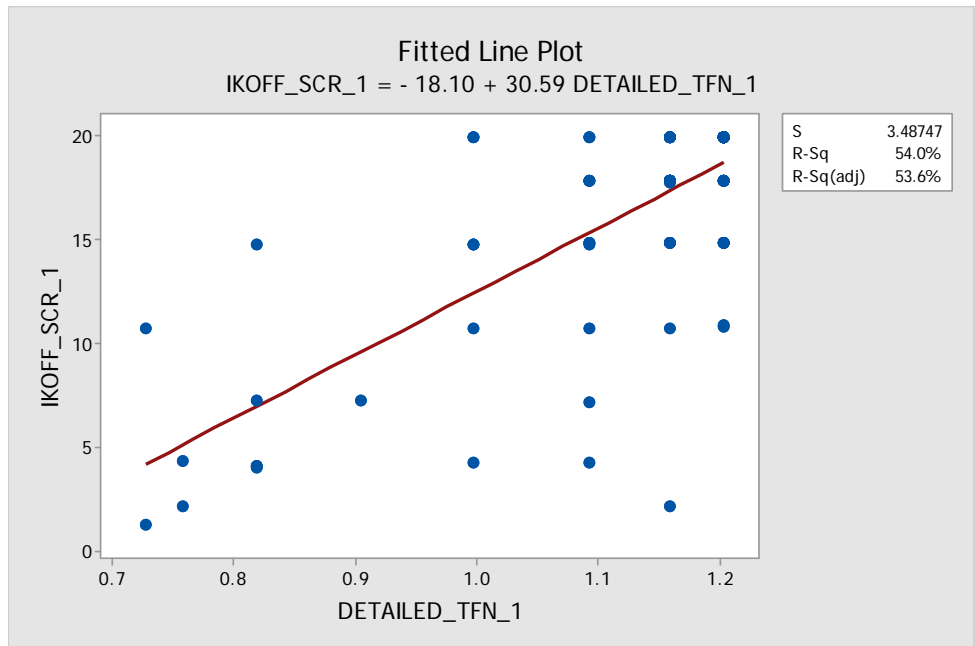
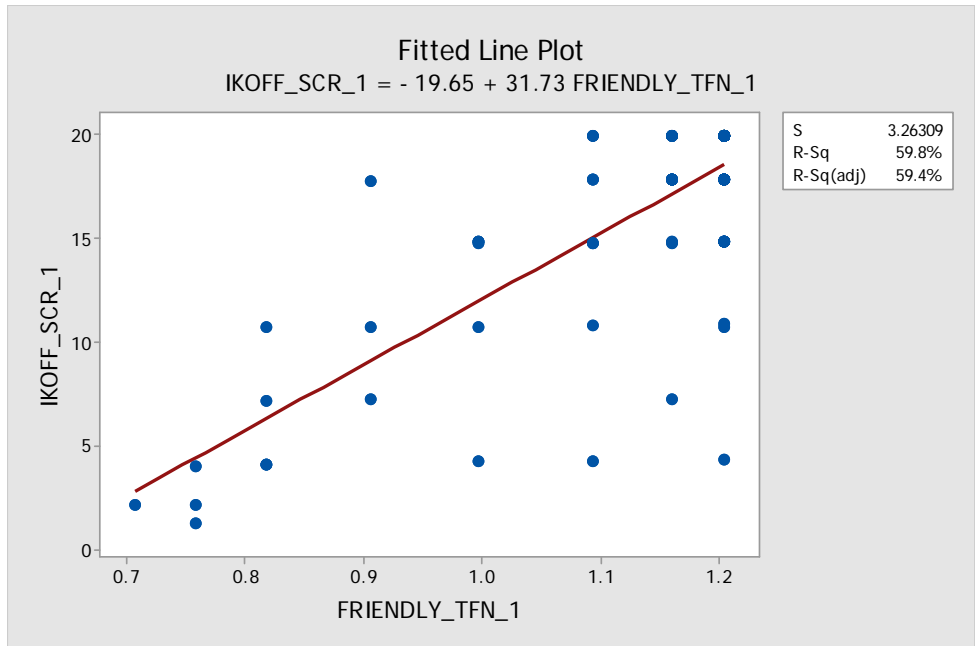
ROWS: FRIENDLY (AMOUNT) Columns: ACADEMIC_CLASS

	GRADS	FRS_SPHS	JNR_SNS	All
1	12	16	15	43
	14.61	13.36	15.03	
2	9	10	6	25
	8.5	7.77	8.74	
3	14	6	15	35
	11.89	10.87	12.23	
All	35	32	36	103

Cell Contents: Count, Expected count.

Now consider how IKOFF relates to a few selected AADQTCs. Employing the 'FITTED LINE PLOT' routine embedded in the Software Package MINITAB 17, we produce the following 4 plots. In all 4 plots there is strong evidence of a linear relation between the transformed AADQTC, PREPARED, INTELLIGENT, FRIENDLY, and DETAILED, respective R²s of 59.8, 71.0, 59.8, and 54.0 %.





Regression Analysis: IKOFF_Y versus INTERESTING_PROFESSOR_X

The regression equation is

$$IKOFF_Y = 1.797 + 5.498 \text{ INTRS_P_X}$$

S = 0.291241 R-Sq = 84.7% R-Sq(adj) = 84.6%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	47.5879	47.5879	561.04	0.000
Error	101	8.5670	0.0848		
Total	102	56.1549			

Regression Analysis: IKOFF_Y versus INTELLIGENT_TFN_1

The regression equation is

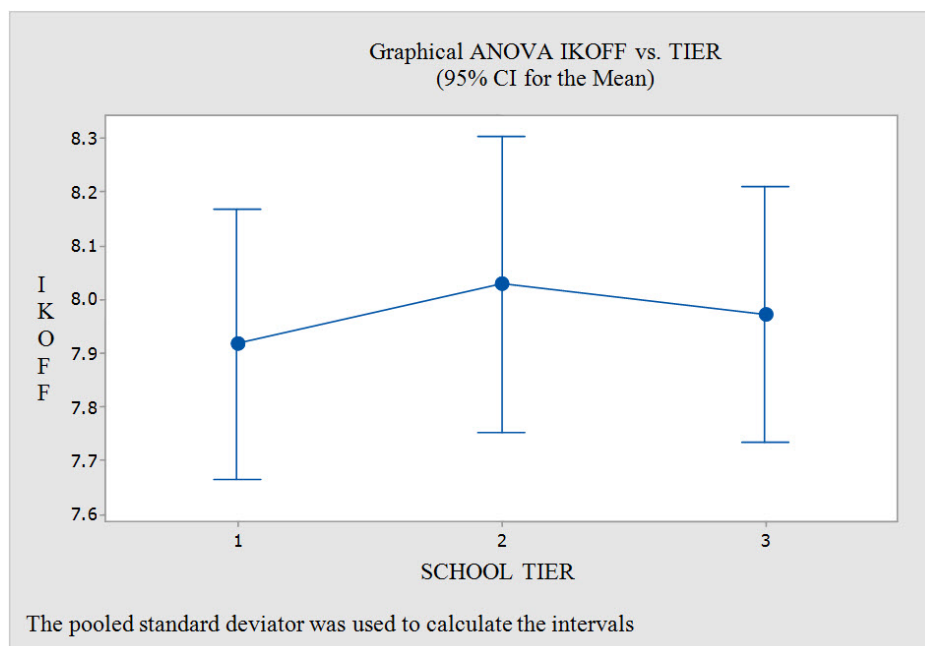
$$\text{IKOFF_Y} = -0.1024 + 6.947 \text{ INTELLIGENT_TFN_1}$$

$$S = 0.280559 \quad R\text{-Sq} = 85.8\% \quad R\text{-Sq(adj)} = 85.7\%$$

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	48.2048	48.2048	612.41	0.000
Error	101	7.9501	0.0787		
Total	102	56.1549			

It seemed interesting to raise the question: do students at large, 'high dollar' (Tier 1) schools have different 'professor and course experiences' than those at Tier 2 and Tier 3 schools?



This study included students from a broad range of universities and colleges categorized according to three tiers: 1[Universities such as Michigan, Alabama, Duke, LSU, West Virginia, Vanderbilt ...]; 2[schools such as Loma Linda, Howard, Troy, Georgia Southern, Tulsa, Suffolk ...]; and 3[schools such as Tuskegee, Hampton, Alabama State, Berea ...]. The above graph captures the results of a ONE-WAY ANOVA. IKOFF scores for the three tiers are not statistically significantly different.

DISCUSSION, CONCLUSION AND FUTURE WORK

Throughout, our approach has been exploratory. The size of our data set has been inhibitive to a significant degree, but that shortcoming aside, what has been demonstrated is that our approach is promising. It appears to open the way for deeper, richer, more powerful, and more enlightening analyses of student appraisal of their academic experiences with respect to courses and professors (teachers). Over the course of the next 10-12 months, we expect to quadruple the number of respondents and carry out analyses suggested by the limited data gathered in this initial effort.

Bibliography

- Aleamoni, L. M. (1999). Student rating myths versus research facts from 1924 to 1998. *Journal of Personnel Evaluation in Education*, 13(2), 153-166.
- Apodaca, P., & Grad, H. (2005). The dimensionality of student ratings of teaching: integration of uni - and multidimensional models. *Studies in Higher Education*, 30(6), 723-748.
- Baker, P. C., & Remmers, H. H. (1952). The measurement of teacher characteristics and the prediction of teaching efficiency in college. *Review of Educational Research*, 224-227.
- Burdsal, C. A., & Bardo, J. W. (1986). Measuring student's perceptions of teaching: Dimensions of evaluation. *Educational and Psychological Measurement*, 46(1), 63-79.
- Cashing, W. E. (1990). Student ratings of teaching: A summary of the research. *Instructor*, 69, 34.
- Cashing, W. E., Downey, R. G., & Sixbury, G. R. (1994). Global and specific ratings of teaching effectiveness and their relation to course objectives: Reply to Marsh (1994).
- Coffman, W. E. (1954). Determining students' concepts of effective teaching from their ratings of instructors. *Journal of Educational Psychology*, 45(5), 277.
- Deshpande, A. S., Webb, S. C., & Marks, E. (1970). Student perceptions of engineering instructor behaviors and their relationships to the evaluation of instructors and courses. *American Educational Research Journal*, 289-305
- Feldman, K. A. (1976). The superior college teacher from the students' view. *Research in Higher Education*, 5(3), 243-288.
- Feldman, K. A. (1988). Effective college teaching from the students' and faculty's view: Matched or mismatched priorities?. *Research in Higher Education*, 28(4), 291-329.
- Feldman, K. A. (1996). Identifying exemplary teaching: Using data from course and teacher evaluations. *New Directions for Teaching and Learning*, 1996(65), 41-50.
- Feldman, K. A. (2007). Identifying Exemplary Teachers and Teaching: Evidence from Student Ratings¹. In *The scholarship of teaching and Learning in higher education: An evidence-based perspective* (pp. 93-143). Springer Netherlands.
- Freeman, H. R. (1994). Student evaluations of college instructors: Effects of type of course taught, instructor gender and gender role, and student gender. *Journal of Educational Psychology*, 86(4), 627.
- French-Lazovik, G. (1974). Predictability of students' evaluations of college teachers from component ratings. *Journal of Educational Psychology*, 66(3), 373.
- Gage, N. L. (1961). The appraisal of college teaching: An analysis of ends and means. *The Journal of Higher Education*, 17-22.
- Hildebrand, M. (1971). *Evaluating University Teaching*.
- Jackson, D. L., Teal, C. R., Raines, S. J., Nansel, T. R., Force, R. C., & Burdsal, C. A. (1999). The dimensions of students' perceptions of teaching effectiveness. *Educational and Psychological Measurement*, 59(4), 580-596.
- Kulik, J. A., & McKeachie, W. J. (1975). The evaluation of teachers in higher education. *Review of research in education*, 210-240.
- McKeachie, W. J. (1979). Student ratings of faculty: A reprise. *Academe*, 65(6), 384-397.
- Marsh, H. W., & Hocevar, D. (1984). The factorial invariance of student evaluations of college teaching. *American Educational Research Journal*, 21(2), 341-366.
- Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International journal of educational research*, 11(3), 253-388.

- Marsh, H. W., & Hocevar, D. (1991). The multidimensionality of students' evaluations of teaching effectiveness: The generality of factor structures across academic discipline, instructor level, and course level. *Teaching and Teacher Education*, 7(1), 9-18.
- Marsh, H. W. (1994). Weighting for the right criteria in the Instructional Development and Effectiveness Assessment (IDEA) system: Global and specific ratings of teaching effectiveness and their relation to course objectives. *Journal of educational psychology*, 86(4), 631.
- Mueller, F. J. (1951). Trends in student ratings of faculty. *Bulletin of the American Association of University Professors*, 319-324.
- Radmacher, S. A., & Martin, D. J. (2001). Identifying significant predictors of student evaluations of faculty through hierarchical regression analysis. *The Journal of Psychology*, 135(3), 259-268.
- Waters, M., Kemp, E., & Pucci, A. (1988). High and low faculty evaluations: Descriptions by students. *Teaching of Psychology*, 15(4), 203-204.
- Wright, R. E. (2006). Student evaluations of faculty: Concerns raised in the literature, and possible solutions. *College Student Journal*, 40(2), 417.