Q1: 2016 Winter Quarter Analysis

Original hypotheses:

H1: Did our intervention(s) impact students attitudes and emotions towards mathematics?
Math Perceptions Survey Potential Question Matches:
1,4,7,12,14,15,16,22,24,25,27,31,36,37

H2: Did our intervention(s) impact students views of the relevance of math in their lives and in society?
Math Perceptions Survey Potential Question Matches: 2,6,9,13,20,28,30,33,35,39

H3: Did our intervention(s) impact how students think about learning math?
Math Perceptions Survey Potential Question Matches: 5,17,21,26,32,38

H4: Did our intervention(s) impact willingness to persevere?
H4a: From Math Perceptions Survey – looking at any increase in perseverance in math
Math Perceptions Survey Potential Question Matches: 3,8,10,11,18,19,23,29,34
H4b: From Grit Survey – looking at any increase in perseverance in general.
Grit Scale is unidimensional

Analysis:

1) Cohesion of the questions involved in each hypothesis:
    a. The questions H1, H2, H4a & H4b had high correlations, meaning they are measuring the same underlying construct (idea) – this is measured using Cronbach's Alpha, which was above .85 for each of those (above .7 is considered good).
    b. H3 Alpha score was initially too low, meaning the items didn't correlate well, and are probably not measuring one single idea. The initial measure was .59. I eliminated Q21 and Q38 from that group. The resulting cohesion measure was .71, so into the "good" realm.

2) Paired t-tests of the difference between Pre and Post test measures.
    a. Using paired t-tests allows us to compare each person's overall pre and post test scores on each of the hypotheses. (We used the mean of the question set.
    b. A p value less than .05 is "significant", meaning the pretest and posttest values are significantly different. Here are the p-values for the hypotheses:
        - H1 .007 (highly significant)
        - H2 .218 (not significant)
        - H3 .039 (significant)
        - H4a .015 (highly significant)
        - H4b .429 (not significant)

3) Finally, there is the direction of the change – and this is unexpected. In every case the averages went DOWN. The scores were higher (showed better emotions, more math relevance, more

positive thinking, more perseverance, more grit) on the pretests. Here are some thoughts about that:

a. The decreases are small – about 1 to 3 points decrease for each hypothesis (over more than ten questions)
b. A second thought is that students may not have taken math for quite a while, and are coming back to class having forgotten their earlier math perceptions. On the first day they are very "gung-ho" about everything. As the course proceeds, they encounter challenges and are reminded about their previous math experiences.
    - Potential protocol change: don't give the pretest right on the first day. Give students time to acclimatize before the pretest and then perhaps measures will go up by the time of the posttest.
    - Potential protocol change: take an intermediate measure about 1/3 to 1/2 way through the class. There might be a decrease from pre to mid, but then an increase in mid to post.
c. It may be that the students are not homogenous. Perhaps those who are coming from a lower level math class have different attitudes from those who are taking their first class. We would need to go into the student records to find what they have taken already.

4) I looked at regression to determine most important determinants of the Posttest scores. The most important one was Pretest in all cases. It was the only significant determinant for H2 and H4. H1 was also determined by their grade in the course; H3 was also determined by Age and the number of previous developmental math classes the student passed.

5) Using k-means cluster analysis (for 2 clusters) with all the demographic variables plus Pretest scores, the clusters seemed to fall mainly based on high/low pretest score. That was true for all the hypotheses. So, the prior development class variables (how many classes, a derived Y/N variable, PriorDevMathPass) and the final grade in class were not segregating the PreTests.

Regression was also checked on the Pretest with all SMS variables PriorDevMath(Y/N) , PriorDevMathPass (Count) current-course-grade, age, sex.
Nothing significant, except H3/Age was .051 (ie not QUITE significant at the .05 level)

Q2: 2016 Spring Quarter Analysis

Quarter-on-Quarter Analysis:

Levene's Test for Equality of Variances shows that there are no significant differences in variance between the two quarters, or between the two classes within a quarter.

For H1 – H3, there are no significant differences between the pre- or post/mid hypotheses scores based on quarter, or between the two class groups within each quarter. Therefore, the two quarters and four classes appear to be mostly homogeneous. For H4 (the "Grit" portion of the Perception survey), both Pre-test and Post tests do differ between Math093 and Math098, which were offered Winter quarter. The significance level for difference between classes for H4_Pre scores is .063 – which, while not significant, bears notice because H4_Post scores are significantly different (.031). This may indicate that there are some differences between students in these classes, or perhaps as they progress through the sequence.

I did not previously check for differences between the classes. There is, however, a trend (not significant) that the 093 classes have the lowest scores, 095 has scores in between the other two classes, and 098 have the highest scores, on both Pre and Post tests.

Adding a Mid-term survey

Unfortunately, the lack of difference between Winter quarter's Post and Spring's Mid scores deny our hypothesis (or hope) that the mid-quarter survey from this term would yield results that were lower than the end-of-quarter survey. The Midterm test was not added for the high school class.

In first analysis, the Spr Q differences between Pre and Mid were NOT significant for any Hx. This finding of non-significant difference between the scores led the data analyst to attempt factor analysis in the hope of finding different, and hopefully more useful, results.

Factor (or Component) Analysis:

Factor Analysis shows which questions are highly correlated to each other, indicating an underlying factor or set of questions for a hypothesis. An attempt was made to find the underlying components of the hypotheses, to see if there were underlying factors in the questions that were different from our original understanding of how questions would match up to hypotheses. Also, since there has been no published testing of the Perceptions survey, we hoped a factor analysis might generate some more stable information.

Questions removed due to correlation issues:

- Low correlation:
    - Q18. I complete assignments outside of class: *Low correlation to any factor*
    - Q21. Mathematics is mainly about having a good memory: *Correlated with only Q29*
    - Q23R "I almost never come to class with finished homework correlated only Q16, 21R
    - Q29. I ask questions when I am unsure of a problem: *Correlated with only Q21*
- High correlation: No questions removed due to high correlation: All r < .84

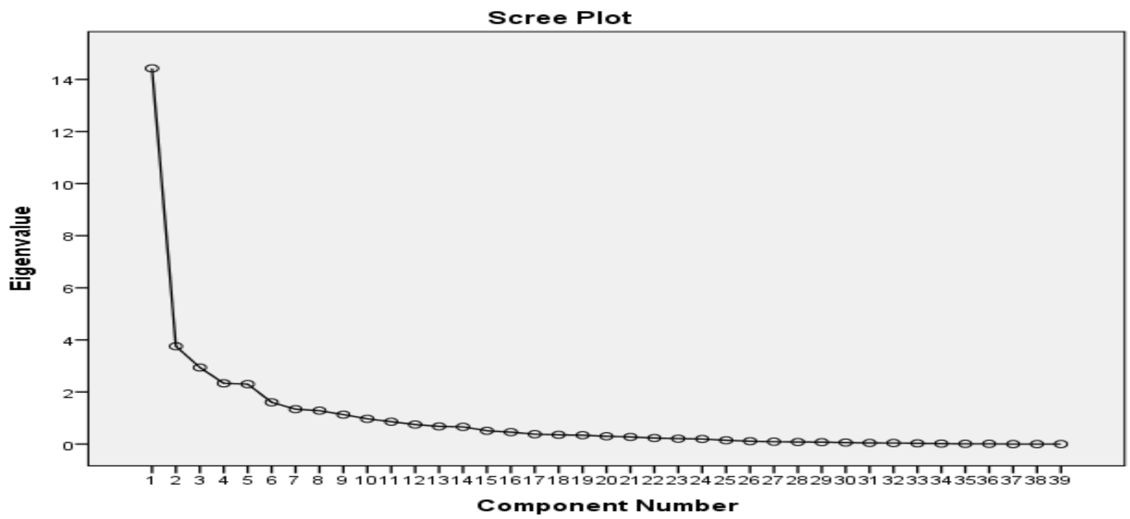
Principal components analysis

Method = PCA
Analyze = Correlation matrix
Extract = Eigenvalue >1.0

- 9 factors were retained originally (eigenvalue >1). Total variance explained: 80.740%
- Looking at the skree plot, 3-4 factors seemed to get the biggest contributions.

In examining the loading on 3 vs. 4 factors, 3 factors made the most logical sense. When trying for 4 factors, the questions did not align well – meaning that questions such as " I work hard in my math classes" and "I give math assignments my best effort" fall into F1 (positive attitude toward math) rather than F3 (try hard). Similarly, the "relevance" questions get split up between F2 (math has relevance) and another factor.

Figure 1: Scree plot of eigenvalues for factor analysis



The H3 questions ("how students think about learning math") factored into both F1 Positive attitude and F3 Trying hard. Ultimately, I could not justify 4 factors. The H3 factor did not "hold together" as an underlying component. Communalities (proportion of each variable's variance that can be explained by the factors) are all above .62 for the remaining three factors.

Many of the questions which were originally hypothesized to go with H4 Grit instead loaded more highly on F1 Positive attitude, while many originally hypothesized to go with H1 Positive attitude loaded more highly on F3 Trying hard. Here is the mapping:

All components had the necessary minimum of three variables contributing to their variance (Velicer & Fava, 1998). There were between 9 and 16 questions per factor.

Table 2: The original 4 hypotheses (Hx) and the 3 Factors generated (Fx)

H1	POSITIVE ATTITUDE/EMOTION
H2	RELEVANCE / IMPORTANCE OF MATH
H3	HOW STUDENTS THINK ABOUT LEARNING MATH
H4	GRIT/PRESERVERENCE

F1	POSITIVE ATTITUDE/LIKE MATH
F2	RELEVANCE / IMPORTANCE OF MATH
F3	TRYING HARD

The variables are not normally distributed (p=.000 for all variables). This means we need to use specific tests to deal with that. However, for factor analysis we need only do KMO (test for sampling adequacy) & Bartlett Test for sphericity (to ensure moderate inter-correlations). Normality test can be over looked.

KMO and Bartlett's Test: Bartlett tests whether there is adequate difference to create separate factors (H0 is that correlation matrix is the Identity matrix, ie perfect correlation), which is rejected at sig. = .000

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.
.854 KMO   > .5 indicates adequate sample size

Bartlett's Test of Sphericity Approx.
Chi-Square = 2370.221
df = 741
Sig. = .000

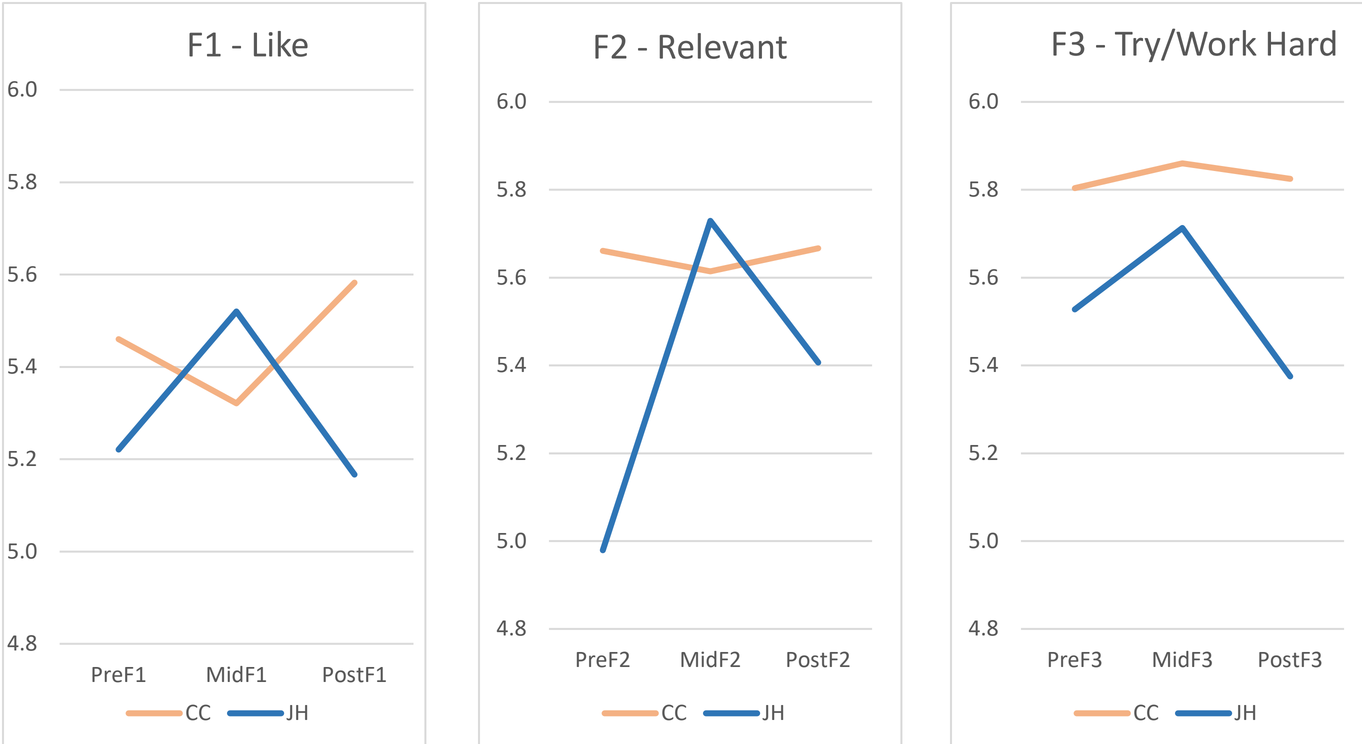
If the assumption of multivariate normality is "severely violated" they recommend one of the principal factor methods; in SPSS this procedure is called "principal axis factors" (Fabrigar et al., 1999). Principal Components factorization assumes that the factors are not correlated (they are orthogonal). That does not seem realistic with this data. Oblimin rotation allows non-ortagonal rotation.

The pretest variables are a highly correlated matrix: over 33% are correlated at .40 or more. The posttest matrix is similarly highly correlated (46% correlations are over .40). The comparison matrix, Pre-test to Post-test is much less correlated (7% of the variables were correlated at .40 or above).

Factors in the Pre, Mid and Post tests

The factor scores in the Pre-test, Mid-Test and Post-test scores show a very interesting pattern in Figure 3. This data shows the average of all students who had any survey scores. We know that many dropped the course, or did not have Pretests. For comparison, Figure 4 has only students with "matched" pre-mid-post tests, meaning one student has taken all three tests. This set may be different from other groups, such as students who dropped the course (therefore have no Post or Mid scores), as well as those who add late (no Pre score and/or no Mid).

Figure 3: Pre, Mid and Post tests for the college classes



Counts for each test/survey:

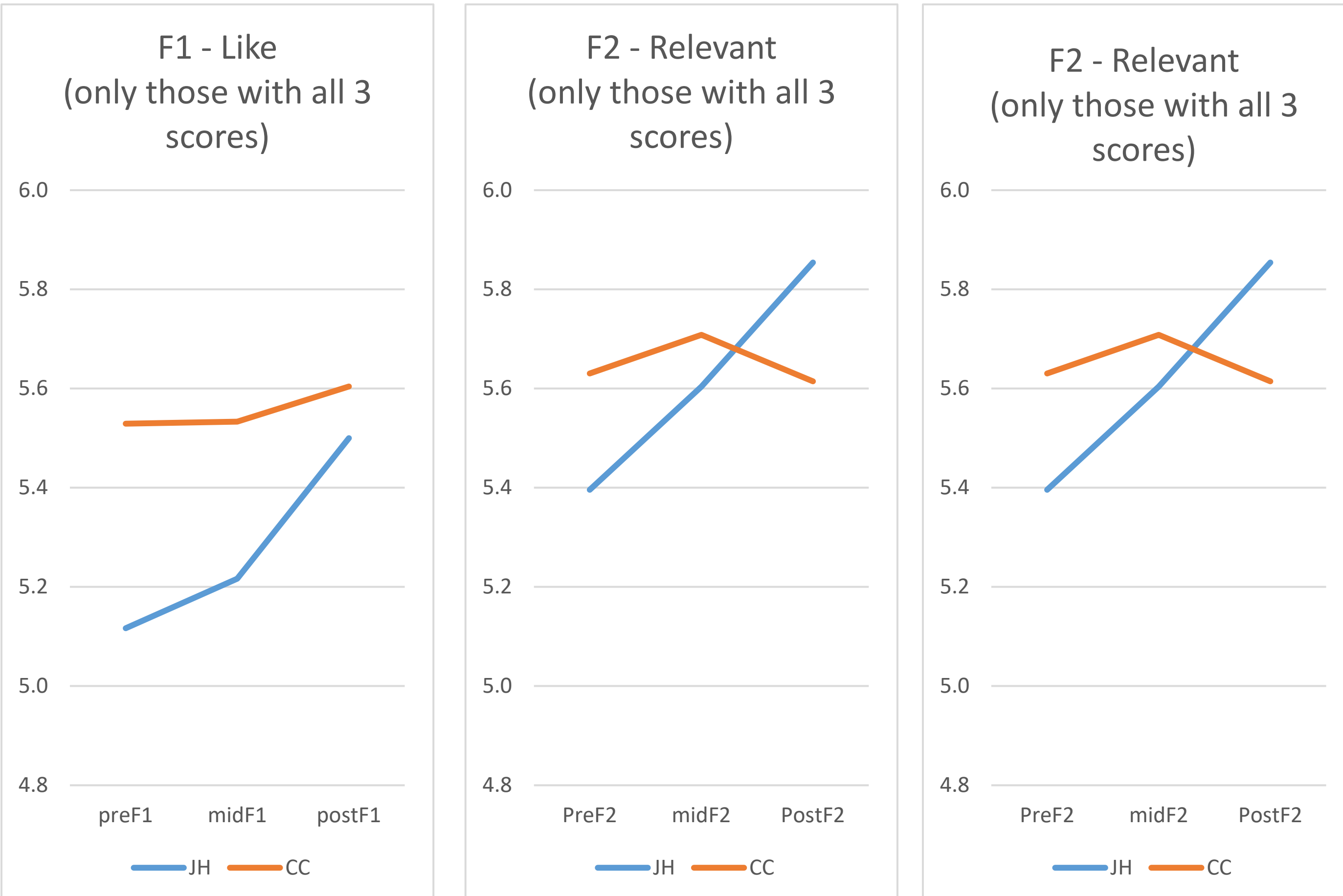
F1	Pre	Mid	Post
JH	16	10	8
CC	32	29	21

F2	Pre	Mid	Post
JH	16	12	8
CC	32	29	20

F3	Pre	Mid	Post
JH	16	12	8
CC	32	29	21

Only the colleges did the midterm survey for spring, since the high school semester was almost over by by first week of May. The midterm spike effect is very interesting, as is the different shape between the two instructors. The instructors teach different classes, as well, which may account for the difference. JH taught 093 and CC taught 094.

Figure 4: Only matched Pre, Mid and Post records (4 for JH and 16 for CC).



It appears people who dropped had higher Pretest scores (the average without them is lower). JH's Midterm "bounce" is also less pronounced – meaning either that those who added late might have had high midterm scores, or that those who dropped had high midterm scores. The Postest scores are quite a bit higher (.2) for those who completed all three surveys, this would imply that those who added late had lower final scores.

Deciding to abandon the Perceptions survey (and factors)

Initially, with the Winter quarter data, we found 3 solid factors, which related quite well to the original Hypotheses H1, H2 and H4. There was some overlap with the original question sets, although some differences too. Several questions did not seem highly correlated with any of the three, however.

Unfortunately, the Spring data did not fit consistently into the same three hypothesized factors. There was substantial overlap but over a third of the individual questions did not line up with the factors as identified in the Winter data. If I looked at both quarters combined, I was back on track with the factors, with only 15% of the questions not lining up with their Winter quarter factors.

Another negative: the Post questions did not align with the suggested Pretest factors. Over 50% did not fall into the previously suggested factors according to the Pretest.

Preliminary work with the factors does not show anything different than what we already reported with the original H1, H2, H3, H4 hypotheses: there is no significant difference in the Pre to Post/Mid tests for the combined classes & quarters. Also no difference between the Post factors Winter Q and the Mid factors Spring Q.

In the end, we decided to abandon the factors, and the entire Perception survey and replace it with the Self-Knowledge survey.

Further spring investigations:

A few questions remained somewhat stable, meaning their pre to post correlation was over .40. These questions seem to be less influenced by what happens in the classroom, such as tests, success or failure, or interventions.

Table 3: Questions which remain largely stable pre-test to posttest

01. Math is something I can learn
03. When I have trouble solving a problem, I try a new strategy
07. I feel accomplished when I solve a problem
08. If I cannot solve a math problem quickly, I quit trying
15. Everyone can learn mathematics
16. I can solve word problems
22. I like doing mathematics
25. Math can be fun
26. Discussing different solutions is a good way of learning math
31. I'm certain I can figure out how to solve difficult math problems
32. Time spent learning why a solution works is time well spent
36. Even if the concepts in math class are hard I can learn them
37. It's important for me to do really well in math

College & High School Comparison:

The sample size for the High School algebra 2 cohort was very small (11). An analysis was undertaken to determine if high school students are different from college students, or which college students they were similar to or different from. The higher up the developmental ladder, the greater the variance between high school student responses and college responses. The sequence in developmental math is 093, 094, 098 (progressing into higher-level concepts).

Figure 4: Growing Attitudinal Variances as Students Progress Up the Ladder of Education

H.S. To Math 093	Comparing High School Algebra students to Math093 students showed only 3 differences in variance at the .05 level.	20. My only interest in math is getting a passing score 21. Mathematics is mainly about having a good memory 34. I give math assignments my best effort
H.S. to Math 093	Comparing High School Algebra students to Math094 students showed 3 additional questions with significant differences in variance at the .05 level.	20. My only interest in math is getting a passing score 21. Mathematics is mainly about having a good memory 34. I give math assignments my best effort 18. I complete assignments outside of class 33. Doing well in math class will help me get a good job when I am done 37. It's important for me to do really well in math
H.S. to Math 098	Comparing High School Algebra students to Math098 students showed another 6 questions with significant differences in variance at the .05 level.	20. My only interest in math is getting a passing score 21. Mathematics is mainly about having a good memory 34. I give math assignments my best effort 18. I complete assignments outside of class 33. Doing well in math class will help me get a good job when I am done 37. It's important for me to do really well in math 02. I think learning math is important for my future 09. Studying mathematics is a waste of time 12. I believe I can get better at math 13. Math has no use outside of school 32. Time spent learning why a solution works is time well spent 35. I can use what I learn in math class in other subjects

Having some differences makes intuitive sense, in that high school may have a more diverse set of students since it is mandatory for everyone 17 or under. One would expect some narrowing of attributes likely takes place in the decision/capacity to go to college. The higher level students may be significantly older than high school students as well, with different perceptions.

Given the small size of the high school group, it is difficult to place much confidence in these differences, due to lack of statistical power. Even when compared with the highest level math, there were less than a third (12 out of 39) of the variables which had variances were significantly different.

The comparisons below examine both winter and spring cohorts as the significance level was consistent. Also noted are the differences in class. Comparison of the means of pre-tests from the college quarters to the high school scores showed significant differences (p=.001 or smaller) on one-third of questions (see table below). In these cases high school students are significantly more negative than college students in their perceptions of math and their abilities at math. This is true in most statements, but only those below are significant differences.

Table 5: Questions with Significant Differences between H.S. & College data

Higher scores = more strongly "positive" statements

01. Math is something I can learn	3.50 H.S. 6.12 Spring 6.33 Winter
04. My math teachers have been unsuccessful with helping me learn math (Reversed)	3.00 H.S. 4.71 Spring 4.74 Winter
08. If I cannot solve a math problem quickly, I quit trying (Reversed)	3.33 H.S. 5.36 Spring 5.26 Winter
09. Studying mathematics is a waste of time (Reversed)	2.67 H.S. 5.71 Spring 6.23 Winter
12. I believe I can get better at math	4.83 H.S. 6.14 Spring 6.30 Winter
13. Math has no use outside of school (Reversed)	2.17 H.S. 5.95 Spring 6.42 Winter
15. Everyone can learn mathematics	3.83 H.S. 5.86 Spring 5.40 Winter
17. Getting the right answer is more important than understanding why the answer works	3.67 H.S. 5.43 Spring 5.56 Winter
18. I complete assignments outside of class	4.33 H.S. 5.88 Spring 5.93 Winter
23. I almost never come to class with finished homework (Reversed)	3.20 H.S. 5.86 Spring 5.95 Winter
27. Only very intelligent students can understand mathematics (Reversed)	1.80 H.S. 5.55 Spring 5.51 Winter

30. Mathematics has no relevance in my life (Reversed)	2.60 H.S. 5.67 Spring 6.05 Winter
38. Ordinarily students cannot understand math, but must memorize the rules (Reversed)	2.80 H.S. 4.64 Spring 5.00 Winter

Interestingly, some scores were higher for the high school students. These questions had higher scores for high schoolers compared to college students, in both Winter and Spring quarters, except for Q35, which was only higher for Winter. While none of these were significantly different, it might provide food for thought.

Table 6: Questions which had high scores in High School

03. When I have trouble solving a problem, I try a new strategy
11. I can tackle a challenging math problem
16. I can solve word problems
28. Achievement and effort in math class are likely to lead to job success later on
29. I ask questions when I am unsure of a problem
31. I'm certain I can figure out how to solve difficult math problems
35. I can use what I learn in math class in other subjects

The table below shows the 10 big "losers" for high school and college, between pretest and posttest. The scores reported are the difference from the Pre to the Post (Pre - Post).The tables are sorted into question number order, and color coded to match the lists below. With only 7 respondents in the H.S. data, we have to take that information with a grain of salt. The red highlighted questions are found in both the college and high school setting, meaning that at both the high school and college level, these lost the most points from Pre- to Post-test. Below these tables are the question text, by factor (color)

Table 7: Biggest Losers: Largest Loss from Pre to Post in College and High School (red shows commonalities)

All college	N Valid	N Missing	Mean	S.D.
PRE_POST_DIFFQ8	72	32	(0.63)	1.84
PRE_POST_DIFFQ17	72	32	(0.46)	1.77
PRE_POST_DIFFQ23	72	32	(0.67)	1.60
PRE_POST_DIFFQ24	71	33	(0.85)	1.69
PRE_POST_DIFFQ25	71	33	(0.45)	1.45
PRE_POST_DIFFQ26	70	34	(0.39)	1.63
PRE_POST_DIFFQ32	71	33	(0.52)	1.50
PRE_POST_DIFFQ33	71	33	(0.45)	1.33
PRE_POST_DIFFQ34	70	34	(0.56)	1.37
PRE_POST_DIFFQ37	70	34	(0.54)	1.38

High school	N Valid	N Missing	Mean	S.D.
PRE_POST_DIFFQ8	7	0	(1.43)	2.23
PRE_POST_DIFFQ16	7	0	(0.71)	1.70
PRE_POST_DIFFQ19	7	0	(0.43)	1.27
PRE_POST_DIFFQ21	7	0	(0.57)	1.27
PRE_POST_DIFFQ23	6	1	(0.67)	0.82
PRE_POST_DIFFQ28	6	1	(0.67)	1.21
PRE_POST_DIFFQ31	6	1	(0.50)	1.05
PRE_POST_DIFFQ34	6	1	(0.50)	1.38
PRE_POST_DIFFQ37	6	1	(1.50)	2.35
PRE_POST_DIFFQ38	6	1	(1.33)	2.50

Factor 1: I have a positive attitude toward math

Q5_1	Math is something I can learn
Q5_3	When I have trouble solving a problem, I try a new strategy
Q5_4R	My math teachers have been unsuccessful with helping me learn math
Q5_8R	If I cannot solve a math problem quickly, I quit trying
Q5_11	I can tackle a challenging math problem
Q5_12	I believe I can get better at math
Q5_14R	I find it difficult to focus during math class
Q5_15	Everyone can learn mathematics
Q5_16	I can solve word problems
Q5_19	I can usually do math problems that take a long time to complete
Q5_22	I like doing mathematics
Q5_24	I think I will do/have done well in mathematics this semester
Q5_25	Math can be fun
Q5_31	I'm certain I can figure out how to solve difficult math problems
Q5_36	Even if the concepts in math class are hard I can learn them

Factor 2: Math is relevant

Q5_2	I think learning math is important for my future
Q5_9R	I think learning math is important for my future
Q5_13R	Studying mathematics is a waste of time
Q5_17R	Math has no use outside of school
Q5_20R	Getting the right answer is more important than understanding why the answer works
Q5_21R	My only interest in math is getting a passing score
Q5_26	Mathematics is mainly about having a good memory
Q5_30R	Discussing different solutions is a good way of learning math
Q5_32	Mathematics has no relevance in my life
Q5_33	Time spent learning why a solution works is time well spent
Q5_35	Doing well in math class will help me get a good job when I am done
Q5_39	I can use what I learn in math class in other subjects

F3: Hard work is rewarding and rewarded

Q5_5	I can work with a partner to find a solution to a problem
Q5_7	I feel accomplished when I solve a problem
Q5_10	I work hard in my math classes
Q5_18	I complete assignments outside of class
Q5_27R	Only very intelligent students can understand mathematics
Q5_28	Achievement and effort in math class are likely to lead to job success later on
Q5_29	I ask questions when I am unsure of a problem
Q5_34	I give math assignments my best effort
Q5_37	It's important for me to do really well in math

Did not fit into factors (highly correlated with more than 1 factor or low correlation with any factor)

Q5_6	I use math in my daily life
Q5_23R	I almost never come to class with finished homework
Q5_38R	Ordinarily students cannot understand math, but must memorize the rules

The table below shows the "winners" for High School and college, between pretest and posttest. These might be areas where you could spend less time and support. The Green highlighted question data show commonalities in college and high school. There were fewer questions which showed a gain for the college students, many for the H.S. students.

Table 8: Biggest Winners: (Greatest gains from Pre to Post test)

All college	N Valid	N Missing	Mean	S.D.
PRE_POST_DIFFQ3	70	34	(0.01)	1.16
PRE_POST_DIFFQ4	72	32	0.24	1.99
PRE_POST_DIFFQ5	71	33	0.07	1.43
PRE_POST_DIFFQ6	70	34	0.60	1.80
PRE_POST_DIFFQ16	71	33	0.00	1.25
PRE_POST_DIFFQ39	71	33	0.18	1.60

High school	N Valid	N Missing	Mean	S.D.
PRE_POST_DIFFQ1	7	0	.71	.951
PRE_POST_DIFFQ3	6	0	.17	1.722
PRE_POST_DIFFQ4	7	0	.86	1.676
PRE_POST_DIFF Q6	7	0	.43	.787
PRE_POST_DIFF Q9	7	0	.57	1.397
PRE_POST_DIFFQ 11	7	0	.29	.951
PRE_POST_DIFF Q13	7	0	.57	1.134
PRE_POST_DIFFQ1 4	7	0	.43	2.507
PRE_POST_DIFFQ 15	7	0	.43	.787
PRE_POST_DIFFQ 17	7	0	.43	1.813
PRE_POST_DIFFQ 20	7	0	.71	.756
PRE_POST_DIFFQ2 4	6	1	.17	.753
PRE_POST_DIFFQ 26	6	1	.33	.816
PRE_POST_DIFF Q29	6	1	.17	.408
PRE_POST_DIFFQ 30	6	1	.17	1.722
PRE_POST_DIFFQ 33	6	1	.33	.816
PRE_POST_DIFFQ 35	6	1	.50	.837
PRE_POST_DIFFQ 39	6	1	.50	.837

Spring Summary

1. The initial Perception survey was difficult to use because the 39 questions did not easily correlate into distinct factors. The researchers initially identified four hypotheses, based on reasonable-assumption groupings, which could be defined as:
    a. H1: Did our intervention(s) impact students attitudes and emotions towards mathematics?
    b. H2: Did our intervention(s) impact students views of the relevance of math in their lives and in society?
    c. H3: Did our intervention(s) impact how students think about learning math?
    d. H4: Did our intervention(s) impact willingness to persevere?

In checking the correlation of the questions involved in each H, there was low cohesion among the questions that had been grouped into the Hs.

Factor analysis was used to create groups of questions with higher cohesion. This resulted in only 3 groups which are described as:

a. F1: positive attitude about/like math
b. F2: math has relevance/ importance
c. F3: I work/try hard at math

Because the factors had higher cohesion than the initial question groups by reasonable-assumption, and because the factors overlay the constructs defined in H1, H2 and H4, the factors were used in the remainder of the spring analysis. However, they are also problematic because the Pre-test Factor group correlations are not similar enough to the Post-test Factor groups.

2. The first priority was to test the H1-H4 (although using the F1, F2, F3). The scores did not change significantly in Pre-to-Post difference measures, and the scores went down rather than up, as hoped. Similarly, there were no significant changes in Pre-to-Mid or Mid-to-Post, although the N on that was small because it was only one quarter.

3. The Factor scores were significantly correlated with each other, and the Pre and Post scores were significantly correlated. The F1Like factor was highly correlated with almost all other scores, Pre, Mid and Post. Pretest F scores were highly negatively correlated with their Posttest counterpart. This means that higher pretest scores fell further. The addition of the Mid-term survey at the college-level added more non-significant results: typically the scores went up at the mid-term, then dropped further at the posttest.

4. A graphic of the Pre, Mid and Post scores separated by instructor showed differences. The level of class is a statistically significant covariate. There were highly significant differences between the "level" of the classes (ie 098 vs. 094 vs. 093/Algebra) and almost all Pre, Mid and Post scores for all Fs and Grit and changes in scores. The only ones not significantly related were the Pre-to-Mid F score changes. One non-significant but interesting difference is that the higher the level of class, the higher the Pre-test scores. Analysis shows that the groups are not homogeneous.

5. The Grit survey was assumed to be a valid construct because it has been previously tested in the literature. It correlated only somewhat with the F3Work factor. The only significant correlations were Pre-Grit to and Post-Grit, and PreF3Work and Post-Grit. There was no significant difference between Pretest Grit score and Posttest Grit score.

6. Covariates were brought into the analysis to see if any of them were related to either the factors or the Grit scores, or the changes in the Pre-to-Post, Pre-to-Mid, and Mid-to-Post scores. Covariates included: Age, Sex, Level of current class, Current grade in class, Number of previous math classes, Avg grade previous math classes, Ethnicity.
    a. Age is positively correlated to all PreF measures, also to PostF1Like and PostF3Work. (Older students like math more, see more significance, see work as producing good outcomes). It was negatively correlated to the change in Pre-to-Post measure for F2Relevant.
    b. Males were significantly more likely than females to see an increase in Pre-to-PostF3Relevance scores

c. Course grade is significantly positively related to all PostF scores. (Better grades meant higher scores in F1Like, F2Relevance, F3Work), also significantly positively related to Increase in F1Like and F2Relevant, but not F3Work.
d. Neither the number of previous math courses, nor the average prior grade in those classes was related significantly to any F scores, pre, post or change.
e. All Pre F scores are significantly correlated. (People who score high on their PreF1Like scores are likely to score the other Pre-scores high. Or low -> low. And so on with all other F measures.)
f. Pre scores are negatively correlated with Post scores. (The higher they started, the lower they fell). This is the case with F1Like, F2Relevant and F3Work.
g. Ethnicity did not show any significant differences, but there were very few non-white students in the group.
h. Grit Pretest scores were significantly related to all covariates (Age, Sex, Ethnicity, Grade in Current Course, Number of prior math classes, and their average grade in those classes. While there was no significant change from PreGrit to PostGrit, an individual's change in Grit score (Post-Pre) was significantly correlated with everything except Gender.

Thoughts moving forward:

- It might be good to change out the Perception survey for a similar survey with higher tested validity, such as ones mentioned in the published literature on attitudes toward remedial math.
- Creating a set of standardized interventions might help surface effects. By that I mean having several defined techniques and try to apply them evenly across all classes at the same times.
    - *This was not realistic, given the different class topics and different students (high school and college).*
- Another thing that would add value to your research would be to have a control group. Perhaps someone in your math faculty that also teaches the 093,094,098 classes would be willing to give the surveys while NOT doing any intervention. It may be that you ARE making a difference, compared to the attitudes of students not in your classes.

Additional Analysis (Summer)

Drop out pattern

I compared completers to drop outs, as far as possible. I used lack of Post scores as indicator of drop out. I know that in some cases the students did not drop out but simply did not complete the Post survey, however that was the best I could do. There were 19 students who were coded as drop outs.

I checked Age, Sex, Ethnicity, Prior Math Classes, and all the Pre scores. Nothing was significantly related to drop out. However, a subsection of the Pre-Grit measures, Q1,6,10 & 12 did predict drop out, for this small group of 19 students.

Interestingly, the very first Grit question ("1. I have overcome setbacks to conquer an important challenge") was quite reliable in predicting drop outs: 95% of those who dropped out answered 1 or 2. What's interesting, though, is that 1 = Very much like me and 2 = Mostly like me. Completers were LEAST likely to agree with the statement, while dropouts were MOST likely to agree. Completers answered 1 or 2 only 30% (somewhat less than random selection, which would be 40%).

Pre-Mid-Post patterns

For only 2 classes did we gather midterm perception data, and some of those students did not complete all three surveys. We had 21 students with Pre-Mid-Post scores (all 3). I was hoping there might be something to learn about how the scores changed and if that went along with other data. Unfortunately, there were several patterns identified, with 3 or fewer in most categories. A cluster analysis on these few students found groupings only on the Perception and Grit scores, not on Age, Sex, Ethnicity, Number of prior math classes or their grade.

Table 9: Changes in Scores –

"Like" Scores:

Pattern from Pre to Mid to Post Scores	Count of students	Avg Pre to Post Grit Drop/Rise	Avg Age	Male: Female: NSP	Avg Grade in class	Avg # Prior Remedial classes
Steady high	3	5.00	20	1:2:0	3.5	.33
Big rise after midterm	3	2.00	25	1:1:1	3.9	0
Steady rise from low	2	-7 & 4	22	0:2:0	3.4	2 for -7 Grit, 0 other
Steady at mid-level	4	-1.50 Range -9 to 4	25 Range 24-31	2:2:0	3.4	.5
Start high, steady decline to average	1	-1	39	F	1.0	0
Pre-Mid downspike	2	-1.50	22	1:0:1	1.7	1
Pre-Mid upspike (low to avg, back to low)	1	6.00	20	F	2.8	0
Steady low-scores	3	-5 Range -12 to 1	19.7	2:1:0	3.8 = +1 Grit Avg 1.6 = Neg	Neg grit had 1.5 avg
Post spike from very low	1	-6	U	U	3.8	0

Note-worthy: * Grit changes – positive grit changes associated with better "Like" patterns, and higher grades
* Prior math classes associated with low grit scores, grades 2.8 and lower

Changes in "Relevant" Scores:

Pattern from Pre to Mid to Post Scores	Count of students	Avg Pre to Post Grit Drop/Rise	Avg Age	Male: Female: NSP	Avg Grade in class	Prior Remedial Math classes
Steady high	4	4.0	23.8	1:3:0	3.5	1 person, had 2.9 grade in class
Big steady rise (*= person with lower Like scores)	2	6* and -7	25* and 20	0:2:0	4.0* and 2.8	0* and 2

Pre-Mid upspike but back to original midlevel	1	-1	21	F	2.2	0
Mid-post upspike	1	6	25	M	4.0	0
Steady at mid-level	4	-5.5	21.0	2:1:0	3.0 (one person=1.1)	Same one had 2 priors
Pre-Mid downspike but back to midlevel	1	2	27	M	3.5	2
Big drop Mid-Post	1	-1	39	F	1.0	0
Steady low (* = one person's scores)	4	1 Range -3 to *7	22.3	2:1:1	2.4 Range 1.0 to *3.7	.75 (* had 0)
Pre-Mid upspike, but back to very low	1	6	20	F	2.8	0
Super low	1	-12	19	F	2.0	1

Noteworthy:

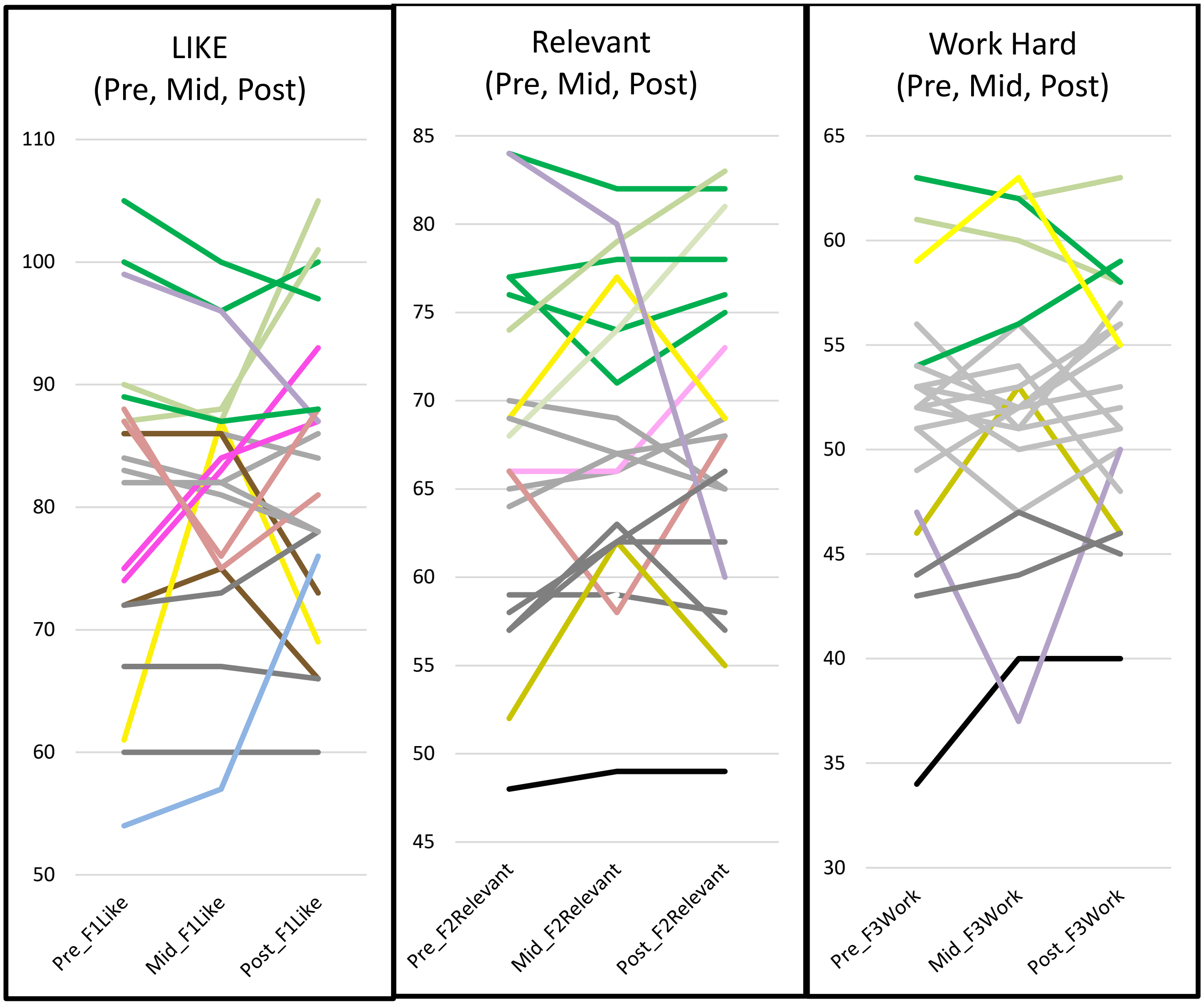
* Grit changes: NOT predictive of Relevant score, positive Grit change associate with high grades
* Average grades in class – C/D grade associated with lower/dropping Relevant scores and negative grit, A/B grades with positive Like and increasing Grit
* Mid-level Relevant scores had B grades, negative grit changes

Changes in "Work Hard" Scores:

Noteworthy: *Grade averages do not track Work Hard scores or Grit changes, although increased grit generally associated with higher grades
* More remedial classes associated with lower grit, lower grades

I tried to look at a student's general pattern of scores (all high, mixed high and low, all low), but that did not correlate with grades, sex, prior math classes, increase or decrease in grit scores. The color coding indicates similar shapes, as outlined in the tables above.

Pattern	**Count of students**	**Avg Pre/Post Grit Drop/Rise**	**Avg Age**	**Male: Female: Unknown**	**Avg Grade in class**	**Count of Prior Remedial Math classes**
Steady high w/ big increases in Grit	2	6	22	0:2:0	4.0	0
Steady high w/ small incr or decrease in Grit	2	.5	21&39	1:1:0	2.0	.5
Steady medium, w/ increase grit & high grades	5	4.6	25.4	2:3:0	3.7	.4
Steady medium, w/ 0 or neg grit change & low grades (*=ONE PERSON)	6	-3.8 *-9	21.2 *19	3:2:1 *F	2.1 *3.2	1.2 *0
Pre-Mid upspike, but back to original (low to mid)	1	6	20.0	F	2.8	0.0
Big midterm drop	1	-6	U	U	3.8	
Steady low	2	-1.0	22	2:0:0	3.8	0
Super low	1	-12	19	F	2.1	1



* Color coding indicates groups, as outlined in tables above

Q3: 2016 Fall Quarter Analysis

New Survey – Self-Knowledge (SK)
We chose to use a new test this term, instead of the Perceptions 39-item test, which was long and the analysis problematic. The Perception items themselves did not seem to be measuring one construct, yet when we tried to divide it logically into factors ("F1: I like math," "F2: I believe math is relevant," "F3: I work hard at math"), these were not stable (individual questions seemed to fall together more or less, depending on whether it was a Pre, Mid or Post Test, and between different quarters.

A new test, the Self-Knowledge (SK) test was substituted. We hoped it would provide a different cohesive score to compare to the Grit test. One of the questions on the SK test, "No one in my family is good at math," does not correlate well with the other measures in the survey, so I removed it from the analysis.

Table 11: Grit Questions (used previously) and Self-Knowledge Questions (new Fall 2016)

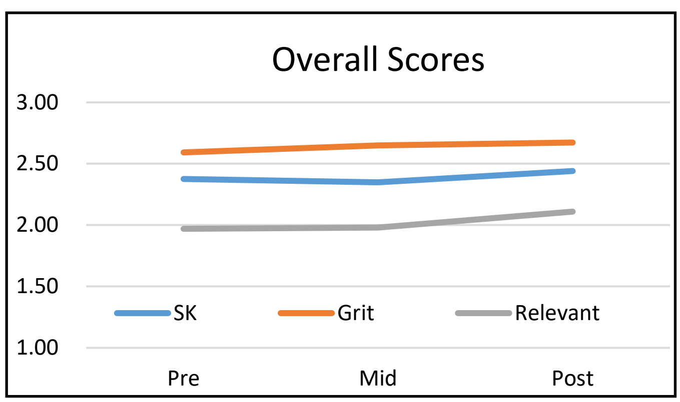
Grit Questions	Self-Knowledge Questions
1. I have overcome setbacks to conquer an important challenge.	1. I find math class very enjoyable.
2. New ideas and projects sometimes distract me from previous ones.*	2. I will never be good at math.*
3. My interests change from year to year.*	~~3. No one in my family is good at math.*~~
4. Setbacks don't discourage me.	4. I have a lot of strategies to use when I tackle a math problem.
5. I have been obsessed with a certain idea or project for a short time but later lost interest.*	5. I get good grades in math classes.
6. I am a hard worker.	6. I feel nervous when I do math problems.*
7. I often set a goal but later choose to pursue a different one.*	7. I give up easily on math problems.*
8. I have difficulty maintaining my focus on projects that take more than a few months to complete.*	8. I think math is a really useful subject.
9. I finish whatever I begin.	9. I want to do well in math class.
10. I have achieved a goal that took years of work.	10. You either get math or you don't.*
11. I become interested in new pursuits every few months.*	*** = Reversed Score**
12. I am diligent.	

In comparing the mean scores of the SK, Grit and Perception Factors they are all highly correlated, yet they have significantly different means. The Perceptions averages were higher, even after normalizing them to a 5 point scale (see next page). Looking at the SK questions as a whole, I would describe them as measuring "I feel good about math." This can be logically associated with "F1: I like math" from our old Perceptions survey and it is highly correlated (.815, where anything over .5 is "strong" and 1.0 is perfect correlation). "F3: I work hard at math" was very similar to Grit, and is significantly correlated. The missing factor, "F2: I think math is relevant," can be associated with one SK question (#8), and indeed they are highly correlated (.718). Interestingly, in this format, the SK Relevant score is a LOT lower than the Perceptions Relevant score. Maybe it's the adjective: "**really** useful subject."

Below are tables for the Pre, Mid and Post tests, Overall and by instructor. Unfortunately, none of these differences are significant. The Perception survey, with its factors, was dropped after the Pre-test, but I am reporting the Pre, Mid & Post SK_Q8 averages for Relevant instead. Pre, Mid and Post tests scores are all significantly different from each other. Again, we see different shapes in the different classes.

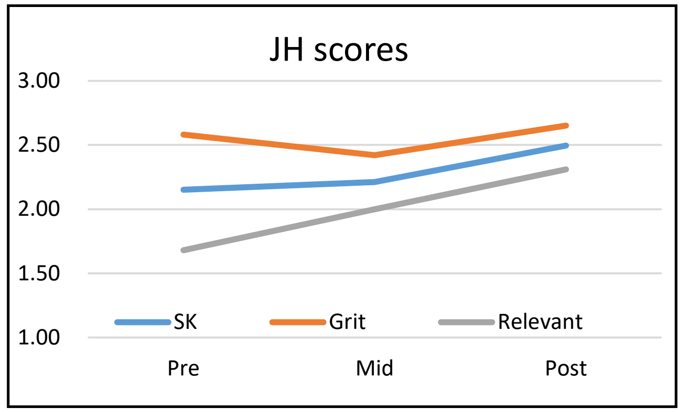
OVERALL Averages - N varies from 73 to 42

	Pre	Mid	Post
SK-Like	2.38	2.35	2.42
Grit	2.59	2.64	2.68
Self_Q8 (relevant)	1.97	2.02	2.17
F1Like	3.68		
F2Relevant	3.78		
F3Work	3.94		



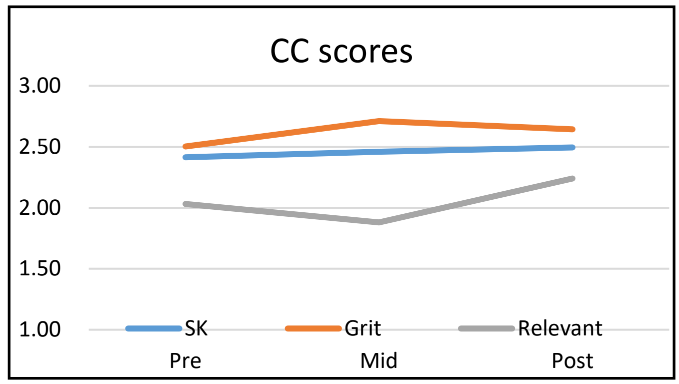
JH Averages N varies from 22 to 13

	Pre	Mid	Post
SK-Like	2.15	2.21	2.45
Grit	2.58	2.42	2.65
Self_Q8 (relevant)	1.68	2.00	2.31
F1Like	3.91		
F2Relevant	3.95		
F3Work	4.16		



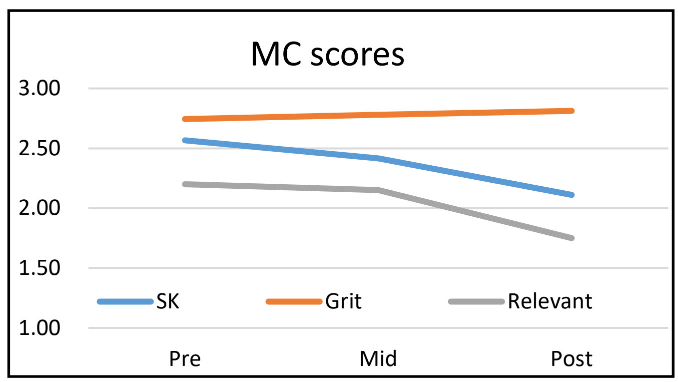
CC Averages N varies from 31 to 16

	Pre	Mid	Post
SK-Like	2.41	2.46	2.49
Grit	2.50	2.71	2.64
Self_Q8 (relevant)	2.03	1.87	2.24
F1Like	3.60		
F2Relevant	3.74		
F3Work	3.98		

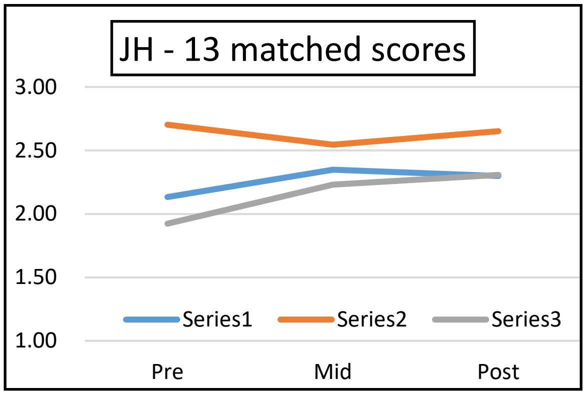


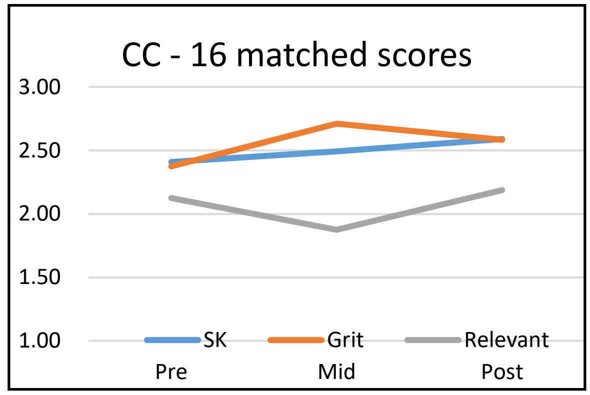
MC Averages N varies from 22 to 8

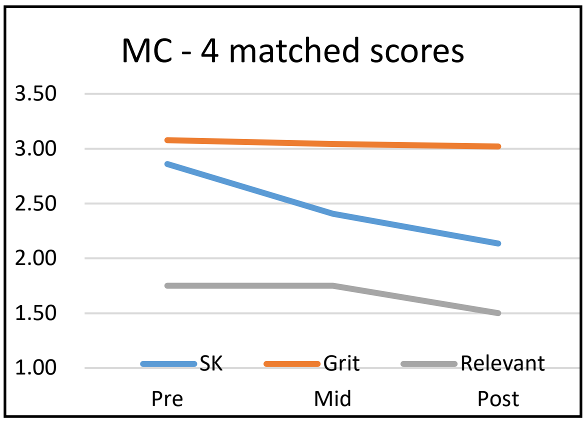
	Pre	Mid	Post
Self Knowledge	2.57	2.42	2.11
Grit	2.74	2.78	2.81
Self_Q8 (relevant)	2.20	2.15	1.75
F1Like	3.57		
F2Relevant	3.67		
F3Work	3.65		



As I did last fall for the spring measures, I wanted to examine only the students where they had all three sets of surveys.







Looking at graphs of individuals' scores, they are all over the place, as before. Some go up and up, some go up and down, some go down and up, some stay flat, some go down and down.

There was no significant effect to any of the students' attitude/belief measures (Grit, Self-Knowledge, or Perception; Pre, Mid or Post) on their final grades.

In measuring Post attitudes, not surprisingly, the Pre-Grit measure is significantly related to the Post-Grit. Oddly, the Pre-SK is not significantly related to the Post-SK.

In trying to look at the big picture, across the terms, we have a problem because we have changed our protocol. We don't have many measures that we have used every term. We have the PrePerception (F1, F2, F3) and the PreGrit, the count of the students' prior developmental math and average grade on those, and their demographic data. If the Sig. value is below .05, then that variable is important. I've highlighted those below.

- Males are significantly more likely to get a higher grade.
- Those who have done well in their prior math classes are significantly more likely to do well now.

a. Dependent Variable: **Grade**

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.408	1.775		1.920	.061
	Inst_Code	-.204	.159	-.167	-1.283	.206
	Coded Level	-.395	.242	-.258	-1.629	.110
	AGE	-.012	.038	-.041	-.331	.742
	Male	.794	.292	.324	2.718	**.009**
	White	-.461	.441	-.119	-1.046	.301
	CountPrior	-.214	.165	-.154	-1.300	.200
	AvgLike_Pre	.269	.249	.192	1.082	.285
	AvgRelevant_Pre	-.109	.271	-.079	-.401	.691
	AvgWork_Pre	-.258	.286	-.141	-.900	.373
	Grit_AvgPre	-.035	.248	-.018	-.139	.890
	AvgGrdPrior	.598	.151	.679	3.965	**.000**

Further analysis

I counted the number of prior math classes and got an average of their GPAs for however many prior classes they had. Regression does not show a specific cutoff, and it could have been one or more classes.

The number of classes were not correlated with their current grade, just the average grade of those classes was correlated. There were 57 people who had prior math classes out of almost 200 students overall. The number or classes:

5 classes: 1
4 classes: 1
3 classes: 6
2 classes: 19
1 class: 30

Looking into that a little more, there doesn't seem to be too strong a relationship near the bottom. For instance,

- Only 5 people had less than 1.0 average in their prior classes. Of those, 2 passed the current class and 3 didn't.
- 6 people who had an average prior grade of less than 2.0 got less than a 2.0 in the current class.
- 7 people who had an average prior grade of less than 2.0 received a 2.0 or better in the current class.
- Only 2 people who got a 3.0 or better in their prior math grade average got less than a 2.0 in in the current class.
- 17 people who got a 3.0 or better in their prior math class got 2.0 or better in in the current class.