Comparison of Freshman Baseline with First Year Seminar and Senior Exiting Assessment Results: Marshall University

Academic Year 2013 – 2014

Summer Assessment Workgroup Members: Harold Blanco, Robert Ellison, Susan Thomas Frank, Marty Laubach, Michael Newsome, Joan St. Germain, Gregg Twietmeyer, Anita Walz, Mary Welch, and Mary Beth Reynolds (Office of Assessment)

Executive Summary

Background

In August 2013, 1,434 incoming freshmen at Marshall University completed baseline assessments that required them to analyze and evaluate information, solve problems, and write effectively. These skills are aligned to three of Marshall University's outcomes; Information Literacy, Inquiry-Based Thinking, and Communication Fluency. In April 2014, 181 graduating seniors completed the same assessments. Freshmen completing Marshall's mandatory First Year Seminar (FYS) completed assessments that were similar to those finished by incoming freshmen and graduating seniors.

In May 2014 a group of nine faculty representing several academic colleges from across the university evaluated a sample of these assessments using a rubric that allowed them to score each assessment across nine criteria (traits). These included *information needed* and *use of information* (Information Literacy), *evidence*, *viewpoints*, and *recommendation/position* (Inquiry-Based Thinking), and *organization*, *genre*, *diction*, and *syntax/mechanics* (Communication Fluency). This project was coordinated by the Office of Assessment.

A random sample of 165 freshman baseline assessments was drawn from the pool of 1,434. To that sample, an additional 59 assessments were added from one of Marshall's academic colleges to assist this college with a specialized accreditation effort. A comparison of the randomly selected sample with the oversample showed no statistically significant differences between the groups on any criterion (trait). Therefore, the additional 59 students were included in our analysis, bringing the baseline sample to 224 (16% of the total number of assessments available). Since only 181 seniors completed the senior exiting assessment, we included all in our analysis, giving us a total of 405 assessments in our sample. One freshman baseline and two senior assessments were discarded because the students did not answer one or more of the prompts. This brought the total number of usable assessments to 402 (223 freshman baseline and 179 senior).

We note that the 181 senior assessment completers did not differ from the exiting senior population (n = 2,345) in terms of race, Honors College enrollment, and entering academic ability. On an ACT scale, mean entering academic ability was 22.8 (population) and 23 (sample). However,

the sample contained a greater proportion of females (68%) than did the population (56%) and the sample's mean college GPA (3.3) was significantly higher than that of the senior exiting population (3.1).

One hundred eighty-eight of the students in the baseline sample of 223 were enrolled in FYS during academic year 2013 – 2014. Of these students, eight did not complete the FYS final assessment, FYS instructors did not return final assessments for 19 of these students to the Assessment Office in time for the assessment team to review them, and 10 students from the baseline sample were enrolled in FYS sections where instructors failed to give students the final assessments. This resulted in a sample of 151 FYS final assessments, results of which we compared to baseline assessments for the same 151 students.

All assessments were de-identified and, for the freshman baseline/senior comparisons, raters did not know which were completed by freshmen and which by seniors. Each assessment had two independent raters. Please see the supporting information that follows this summary for a detailed explanation of scoring procedures.

Comparison of Freshman Baseline to Senior Exiting Results and to Results at the End of FYS

Mean scores (on a scale of 1 – 4) for seniors were significantly higher than freshman baseline measures on all criteria (traits). However, mean performance for seniors ranged from a low of 2.0 (Inquiry-Based Thinking: *recommendations*) to a high of 2.8 (Communication Fluency: *organization*), indicating that there is room for improvement among Marshall's graduating seniors. Mean differences between freshman baseline performance and senior exiting performance ranged from a low of 0.4 for Information Literacy and Communication Fluency (*syntax/mechanics*) to a high of 0.7 Communication Fluency (*genre*).

For the 151 students who completed both baseline and FYS assessments, *paired-samples t-tests* using adjusted alpha levels to control for Type I error (.025 for information literacy), (.017 for Inquiry-Based Thinking), and (.0125 for Communication Fluency) showed a significant mean difference between freshman baseline and FYS results for all criteria (traits) *except* for *information needed*, *viewpoints*, and *syntax/mechanics*. Students showed the largest improvement in performance (0.7) on *recommendation/position*, which aligns with the University's Inquiry-Based Thinking outcome. We note that FYS includes Inquiry-Based Thinking, but not Communication Fluency, among its course outcomes. FYS students showed the greatest gains in evaluating evidence to make a recommendation that acknowledged several sides of an issue. However, students did not demonstrate significant gains in questioning the viewpoints expressed in the pieces of evidence they examined, nor did they make significant gains in indicating the types of additional evidence they might need to make a recommendation. Therefore, we recommend that the FYS Director and course instructors place additional emphasis on helping students to determine information need and critically examine various viewpoints surrounding real-world problems.

Recommendations from the Summer Assessment Workgroup

Assessments

Currently, students are given 90 minutes to complete the freshman baseline and senior assessments. The FYS final assessment, which consists of one additional section asking students to evaluate each piece of evidence that might help them to arrive at a recommendation for accuracy, bias, and relevance, is not included in the freshman baseline/senior assessments. Therefore, students are given 120 minutes to complete the FYS final assessment. The committee recommended standardization among all three assessments (freshman baseline, FYS, and senior) along the following parameters:

- Length of time period to complete – either 90 or 120 minutes
- All sections of assessment present for all
- All presented in electronic format (currently, all assessments are completed using paper and pen/pencil).

The committee recommended implementing the problem-based senior assessment in
- Capstone classes or
- As part of a student's graduation requirements.

Rubrics

Inquiry-Based Thinking (*recommendation/position*): Several reviewers noted this trait as problematic and their reactions were corroborated by poor inter-rater agreement kappa levels (.246 for baseline/senior assessments and .047 for FYS exams). The following changes were recommended:

Rubric Used this session:

Recommendation/Position	• Does not make a recommendation. • Recommendation does not follow assignment instructions. • Does not use evidence in DL to arrive at a recommendation.	Recommendation acknowledges different sides of the issue.	Recommendation takes into account the complexities of the issue.	Recommendation takes into account the complexities of the issue. Any limits to the recommendation are acknowledged.

Recommended rubric for future use:

Recommendation/Position	Either does not make a recommendation or makes a recommendation, but does not justify it in any way.	Recommendation does not acknowledge different sides of the issue.	Recommendation takes into account different sides/complexities of the issue.	Recommendation takes into account the complexities of the issue. Any limits to the recommendation are

			acknowledged.

Communication Fluency (*organization*): Several reviewers noted this trait as problematic and their reactions were corroborated by poor inter-rater agreement kappa levels (.161 for baseline/senior assessments and .213 for FYS exams). The following changes were recommended:

Rubric Used this session:

Organization	Produces a document that is confusing and disjointed.	Produces document that lacks both cohesion and a progression of ideas.	Produces a cohesive document that lacks a clear progression of ideas.	Produces a cohesive document with a clear progression of ideas.

Recommended rubric for future use:

Development	Shows little or no evidence of developing his/her ideas.	Shows some development of ideas.	Shows a strong, but perhaps somewhat incomplete, development of ideas.	Produces a well-organized document in which the ideas have been fully developed.

Communication Fluency (*genre*): Discussion included that this trait was problematic for a couple of reasons. While some reviewers felt that the assignment instructions concerning the genre of writing expected (in this case a memorandum) were not explicit enough, others argued that making the instructions too explicit would reduce this trait to seeing whether or not students could follow a model. They argued that doing so would result in a scoring dichotomy for this trait, i.e. the student either writes a memorandum using correct form or does not. Some suggestions included adding "audience" to the trait name, resulting in *genre/audience* or perhaps calling it *conventions of writing*. Additional feedback from reviewers suggested this revision:

Rubric Used this session:

Genre	Makes no attempt to reproduce specified document format requested in the scenario.	Attempts to produce a document in the specified genre, but the document does not appear to be geared to a specific audience and/or it has formatting errors.	Produces a document in the specified genre, but the document either does not appear to be geared to a specific audience or it has minor formatting errors.	Produces a professional document in the specified genre. Document clearly shows that the writer is speaking to his/her intended audience.

Recommended rubric for future use (note: this language is adapted from the AAC&U Value Rubric for Written Communication):

Genre	Demonstrates minimal attention to basic organization and presentation.	Demonstrates some awareness of basic organization, content, and presentation.	Demonstrates consistent use of important conventions particular to a specific writing task, including organization,	Demonstrates detailed attention to and successful execution of a wide range of conventions particular to a

			content, presentation, and stylistic choices.	specific writing task including organization, content, presentation, formatting, and stylistic choices.

Diction: A suggestion was made to merge this with "syntax" and to rename the trait "Communication Style." The following changes are suggested:

Rubric Used this session:

Syntax/Mechanics	Communication has numerous errors in style, mechanics, or other issues that detract from the message.	Communication has only a few (but noticeable) errors in style, mechanics, or other issues that might detract from the message.	Communication has only a few errors in style, mechanics, or other issues that do not detract from the message.	Uses complex and varied sentence styles or concepts. Communication is virtually free of mechanical, stylistic or other issues that might detract from the message.

Recommended rubric for future use (note: this language comes directly from the AAC&U Value Rubric for Written Communication):

Communication Style	Uses language that sometimes impedes meaning because of errors in usage/mechanics.	Uses language that generally conveys meaning to readers with clarity, although writing may include some errors.	Uses straightforward language that generally conveys meaning to readers. The language in the portfolio has few errors.	Uses graceful language that skillfully communicates meaning to readers with clarity and fluency, and is virtually error-free.

If changes are adopted, our final rubric will be:

Outcomes	Traits	Performance Levels			
		1	2	3	4
Information Literacy	Information Needed	Does not acknowledge or assess the need for more information.	Acknowledges the need for more information but does not identify feasible research methods/sources that would address unanswered questions.	Assesses the need for more information and recommends general research methods/sources that would address some unanswered questions.	Assesses the need for more information and recommends specific research methods/sources that would address most unanswered questions.
	Use of Information	Fails to acknowledge sources from the DL in the final memo.	Indirectly/vaguely acknowledges sources of information from the DL in the final memo.	Clearly acknowledges relevant sources of information from the DL in the final memo.	Integrates relevant information from the DL. Acknowledges sources used in the final memo.
Inquiry-Based	Evidence	Disregards most relevant	Insufficient evidence is taken	Evidence is taken from	Evidence is taken from

Thinking		evidence from the DL.	from sources in the DL or evidence is used without appropriate interpretation/evaluation.	relevant and valid sources in the DL with some interpretation/evaluation, but not enough to develop a coherent analysis or synthesis.	relevant and valid sources in the DL with enough interpretation/evaluation to develop a coherent analysis or synthesis.
	Viewpoints	Ignores most viewpoints expressed in the DL.	Viewpoints expressed in the DL are taken as mostly fact, with little or no question.	Questions some viewpoints expressed in the DL.	Thoroughly questions and evaluates viewpoints expressed in the DL.
	Recommendation/Position	Either does not make a recommendation or makes a recommendation, but does not justify it in any way.	Recommendation does not acknowledge different sides of the issue.	Recommendation takes into account different sides/complexities of the issue.	Recommendation takes into account the complexities of the issue. Any limits to the recommendation are acknowledged.
Communication Fluency	Development	Shows little or no evidence of developing his/her ideas.	Shows some development of ideas.	Shows a strong, but perhaps somewhat incomplete, development of ideas.	Produces a well-organized document in which the ideas have been fully developed.
	Genre	Demonstrates minimal attention to basic organization and presentation.	Demonstrates some awareness of basic organization, content, and presentation.	Demonstrates consistent use of important conventions particular to a specific writing task, including organization, content, presentation, and stylistic choices.	Demonstrates detailed attention to and successful execution of a wide range of conventions particular to a specific writing task including organization, content, presentation, formatting, and stylistic choices.
	Communication Style	Uses language that sometimes impedes meaning because of errors in usage/mechanics.	Uses language that generally conveys meaning to readers with clarity, although writing may include some errors.	Uses straightforward language that generally conveys meaning to readers. The language in the portfolio has few errors.	Uses graceful language that skillfully communicates meaning to readers with clarity and fluency, and is virtually error-free.

Supporting Documentation

Comparison of Freshman Baseline and Senior Exiting Assessment Results

Academic Year 2013 – 2014

Includes Freshman Oversample

Review Procedures

- Each assessment had two independent raters and scores were determined in the following manner:

    1. If raters assigned the same score, that became the score for the artifact.
    2. If raters' scores differed by one point or less, e.g. Rater 1 assigned a score of 1 and Rater 2 a score of 2, the final score was the mean, i.e. 1.5.
    3. If raters' scores differed by more than one point, e.g. Rater 1 assigned a score of 1 and Rater 2 a score of 3, the raters met to discuss the rationale for their scores to see if they could agree on a score or, at minimum, scores that differed by no more than one point.
    4. If raters' scores differed by more than one point and, after discussion, they were not able to resolve the differences, a third rater was assigned to review the assessment.

Rules for Arriving at Final Scores when there were Three Raters: These rules were followed for all assessments conducted.

1. If the third rater's score agreed with one of the first two, the score with the two agreements was used.

2. If the first two raters' scores were two points apart, e.g. 1 and 3 and the third rater's score was in the middle, e.g. 2, the third rater's score was used.

3. If the first two raters' scores were two points apart, e.g. 1 and 3, and the third rater's score was between them, but a decimal, e.g. 1.5 or 2.5, the third rater's score was used.

4. If the first two raters' scores were two points apart, e.g. 1 and 3, and the third rater's score was a "4", the two scores closer together were averaged, e.g. 3.5.

5. IF the first two raters' scores were three points apart, e.g. 1 and 4, the third rater's score was averaged with the closest other rater; e.g. if the third rater's score was 3, the final score was 3.5; if the third rater's score was 2, the final score was 1.5.

Rubric Used for Scoring

Rubric for Freshman/Senior Assessments – Summer 2014

Traits		Performance Levels			
		1	**2**	**3**	**4**
Information Literacy	**Information Needed**	Does not acknowledge or assess the need for more information.	Acknowledges the need for more information but does not identify feasible research methods/sources that would address unanswered questions.	Assesses the need for more information and recommends general research methods/sources that would address some unanswered questions.	Assesses the need for more information and recommends specific research methods/sources that would address most unanswered questions.
	Use of Information	Fails to acknowledge sources from the DL in the final memo.	Indirectly/vaguely acknowledges sources of information from the DL in the final memo.	Clearly acknowledges relevant sources of information from the DL in the final memo.	Integrates relevant information from the DL. Acknowledges sources used in the final memo.
Evidence		Disregards most relevant evidence from the DL.	Insufficient evidence is taken from sources in the DL or evidence is used without appropriate interpretation/evaluation.	Evidence is taken from relevant and valid sources in the DL with some interpretation/evaluation, but not enough to develop a coherent analysis or synthesis.	Evidence is taken from relevant and valid sources in the DL with enough interpretation/evaluation to develop a coherent analysis or synthesis.
Viewpoints		Ignores most viewpoints expressed in the DL.	Viewpoints expressed in the DL are taken as mostly fact, with little or no question.	Questions some viewpoints expressed in the DL.	Thoroughly questions and evaluates viewpoints expressed in the DL.
Recommendation/Position		• Does not make a recommendation. • Recommendation does not follow assignment instructions. • Does not use evidence in DL to arrive at a recommendation.	Recommendation acknowledges different sides of the issue.	Recommendation takes into account the complexities of the issue.	Recommendation takes into account the complexities of the issue. Any limits to the recommendation are acknowledged.
Organization		Produces a document that is confusing and disjointed.	Produces document that lacks both cohesion and a progression of ideas.	Produces a cohesive document that lacks a clear progression of ideas.	Produces a cohesive document with a clear progression of ideas.
Genre		Makes no attempt to reproduce specified document format requested in the scenario.	Attempts to produce a document in the specified genre, but the document does not appear to be geared to a specific audience and/or it has formatting errors.	Produces a document in the specified genre, but the document either does not appear to be geared to a specific audience or it has minor formatting errors.	Produces a professional document in the specified genre. Document clearly shows that the writer is speaking to his/her intended audience.
Diction		Chooses rudimentary/inappropriate vocabulary that confuses the intended meaning of his/her communication.	Chooses vocabulary that conveys the intended meaning of his/her communication.	With the audience in mind, chooses a varied vocabulary that conveys the intended meaning of the communication.	Chooses appropriate, meaningful, clear, and persuasive vocabulary, skillfully communicating meaning to the audience.
Syntax/Mechanics		Communication has numerous errors in style, mechanics, or other issues that detract from the message.	Communication has only a few (but noticeable) errors in style, mechanics, or other issues that might detract from the message.	Communication has only a few errors in style, mechanics, or other issues that do not detract from the message.	Uses complex and varied sentence styles or concepts. Communication is virtually free of mechanical, stylistic or other issues that might detract from the message.

Freshman Baseline/Senior Exiting Comparisons
Mean Scores on a scale of 1 – 4, with 4 being the highest possible score
Freshman n = 223; Senior n = 179
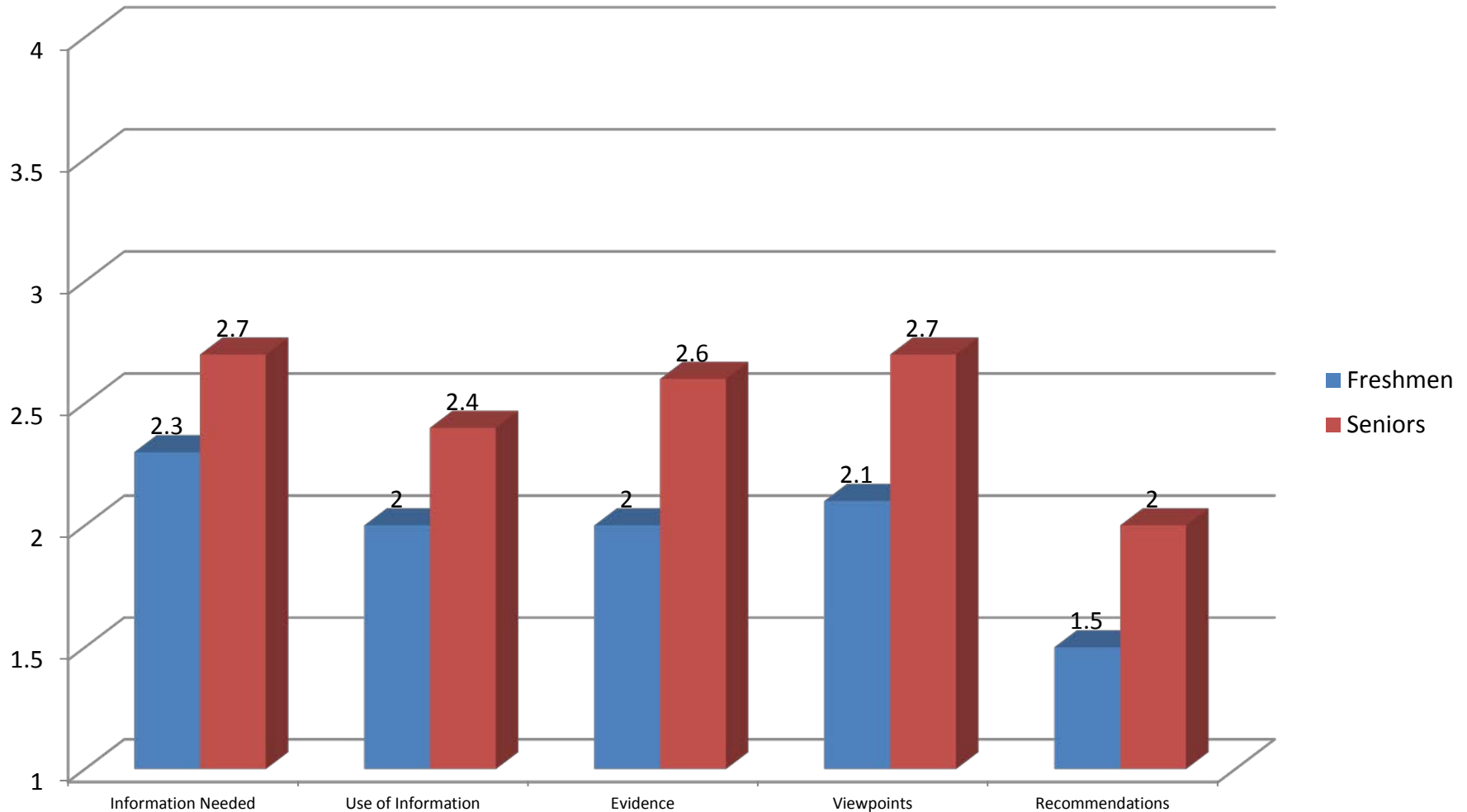
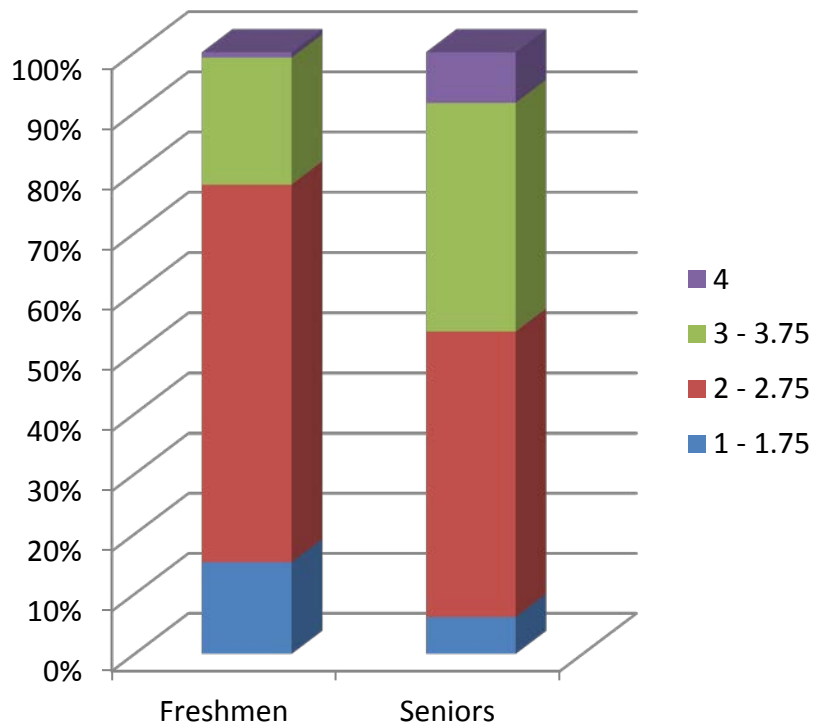# Freshman Baseline/Senior Exiting Comparisons
Freshman n = 223; Senior n = 179

Trait/ Performance Level	Info Needed	Use of Info	Evidence	Viewpoints	Recommendations
1 – 1.75 Freshmen	34 (15%)	101 (45%)	77 (35%)	67 (30%)	161 (72%)
1 – 1.75 Seniors	11 (6%)	46 (26%)	30 (17%)	29 (16%)	87 (49%)
2 – 2.75 Freshmen	140 (63%)	84 (38%)	122 (55%)	116 (52%)	55 (25%)
2 – 2.75 Seniors	85 (47%)	68 (38%)	68 (38%)	57 (32%)	48 (27%)
3 – 3.75 Freshmen	47 (21%)	33 (15%)	24 (11%)	39 (17%)	7 (3%)
3 – 3.75 Seniors	68 (38%)	52 (29%)	63 (35%)	76 (42%)	41 (23%)
4 Freshmen	2 (1%)	5 (2%)	0 (0%)	1 (0%)	0 (0%)
4 Seniors	15 (8%)	13 (7%)	18 (10%)	17 (9%)	3 (2%)
Grand Total Freshmen	223 (100%)	223 (100%)	223 (100%)	223 (100%)	223 (100%)
Grand Total Seniors	179 (100%)	179 (100%)	179 (100%)	179 (100%)	179 (100%)

Freshman Baseline/Senior Exiting Comparisons
Freshman *n* = 223; Senior *n* = 179
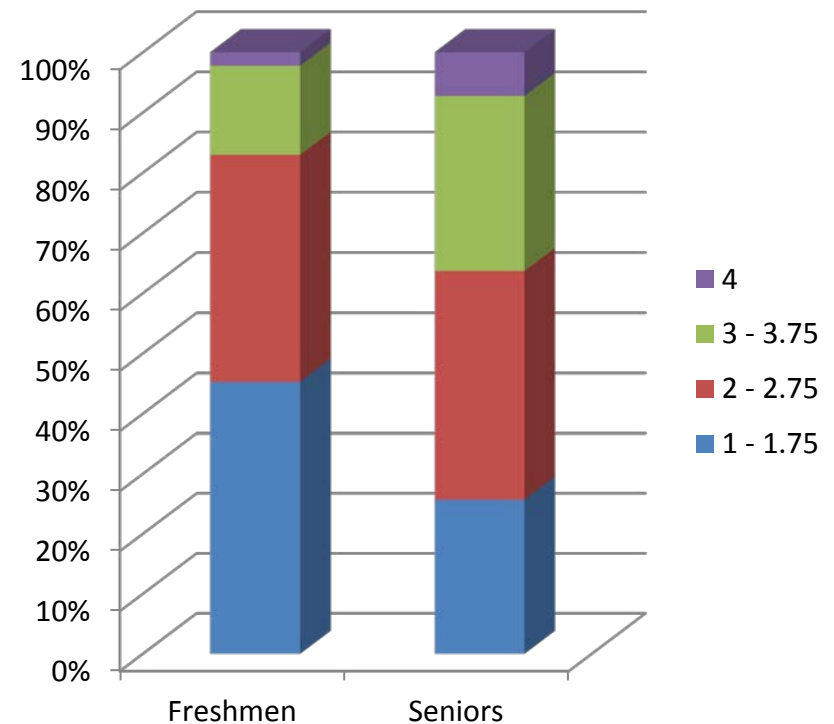
Freshman Baseline/Senior Exiting Comparisons
Freshman *n* = 223; Senior *n* = 179
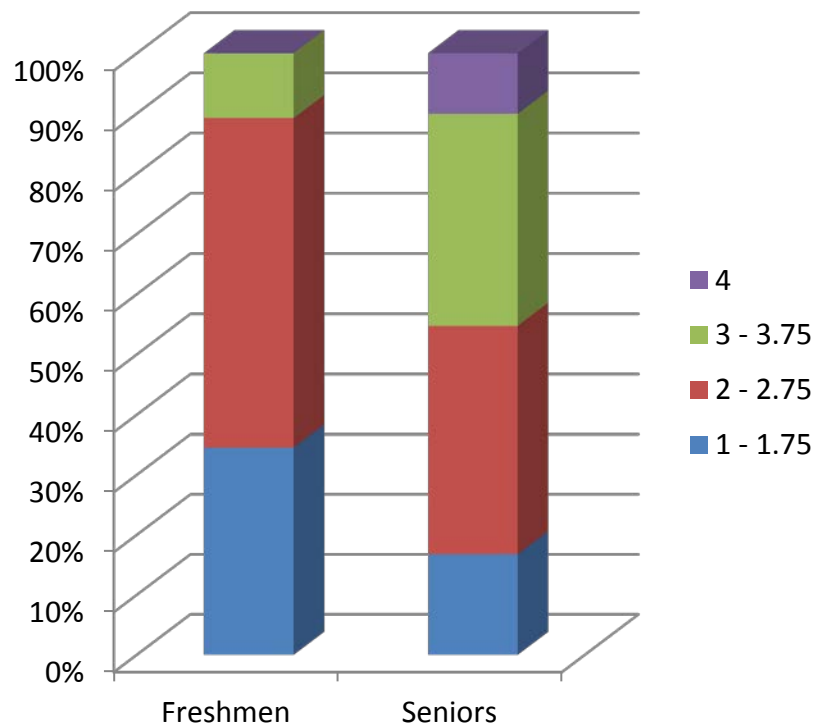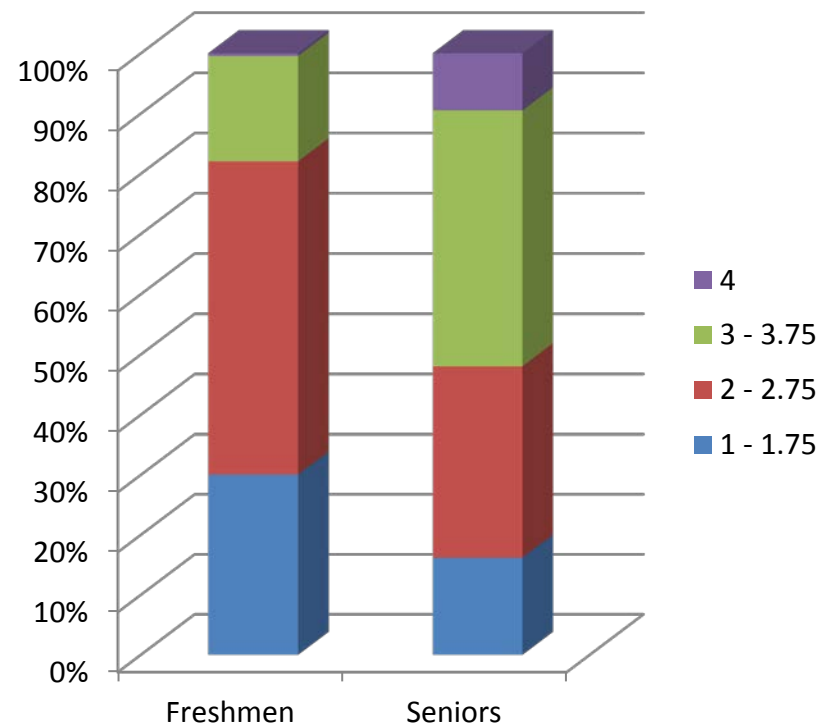
Freshman Baseline/Senior Exiting Comparisons
Freshman *n* = 223; Senior *n* = 179
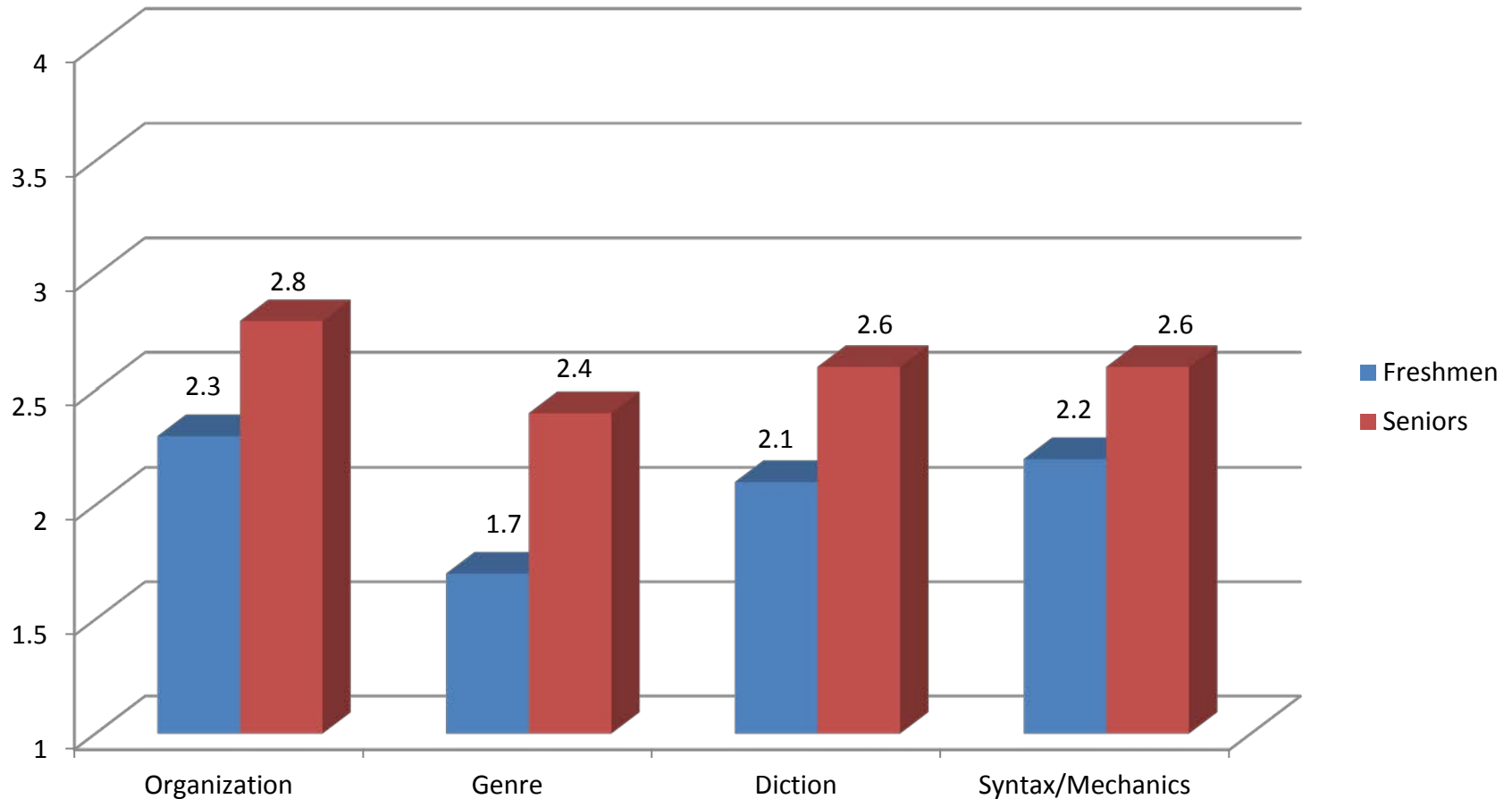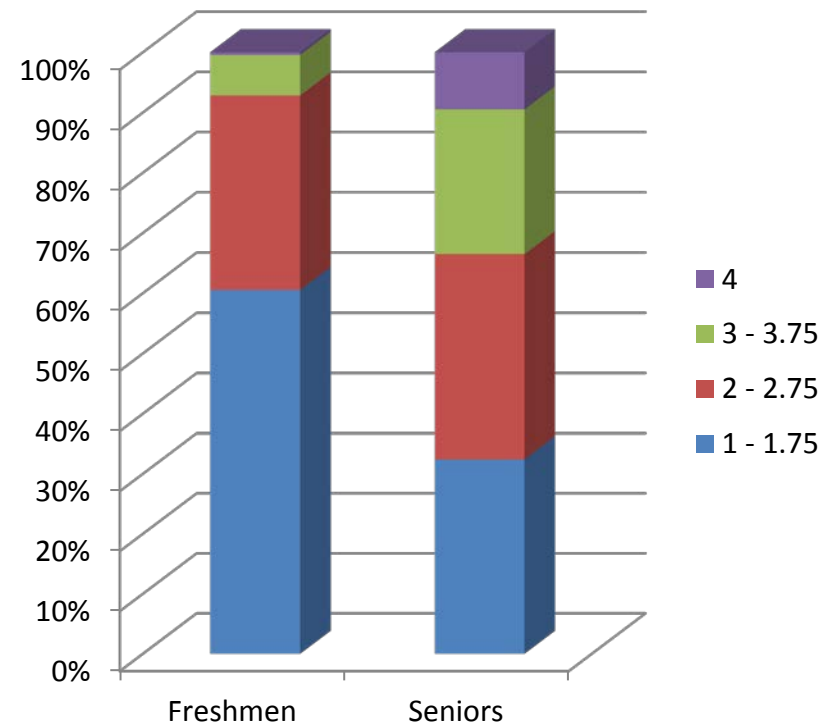
Recommendations
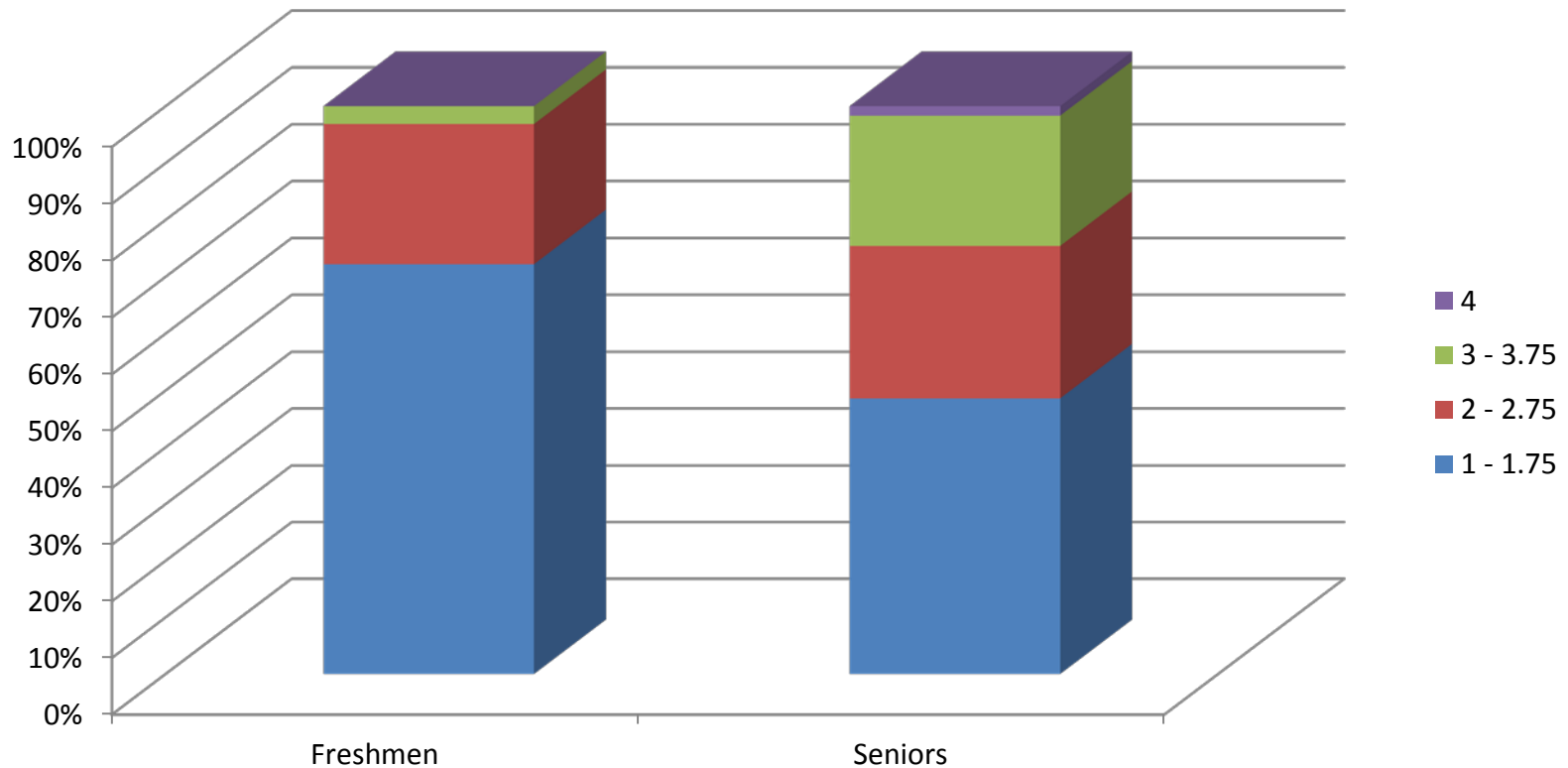
Freshman Baseline/Senior Exiting Assessment Inter-Rater Agreement Results

Trait/ Agreement	Info Needed (Kappa = .388)	Use of Info (Kappa = .333)	Evidence (Kappa = .282)	Viewpoints (Kappa = .350)	Recommendations (Kappa = .246)
Agree	214 (53%)	240 (60%)	200 (50%)	218 (54%)	205 (51%)
Difference = 1 point or less	167 (42%)	152 (38%)	179 (45%)	156 (39%)	154 (38%)
Difference = 1.5 to 2 points	20 (5%)	9 (2%)	21 (5%)	27 (7%)	37 (9%)
Difference = 2.5 to 3 points	1 (0%)	1 (0%)	2 (0%)	1 (0%)	6 (1%)
Total	402 (100%)	402 (100%)	402 (100%)	402 (100%)	402 (100%)

Freshman Baseline/Senior Exiting Comparisons
Mean Scores on a scale of 1 – 4, with 4 being the highest possible score
Freshman n = 223; Senior n = 179

Legend: ■ Freshmen ■ Seniors

Category	Freshmen	Seniors
Organization	2.3	2.8
Genre	1.7	2.4
Diction	2.1	2.6
Syntax/Mechanics	2.2	2.6

Freshman Baseline/Senior Exiting Comparisons
Freshman *n* = 223; Senior *n* = 179

Trait/ Performance Level	Organization	Genre	Diction	Syntax/Mechanics
1 – 1.75 Freshmen	65 (29%)	135 (61%)	55 (25%)	64 (29%)
1 – 1.75 Seniors	22 (12%)	58 (32%)	20 (11%)	28 (16%)
2 – 2.75 Freshmen	108 (48%)	72 (32%)	145 (65%)	113 (51%)
2 – 2.75 Seniors	59 (33%)	61 (34%)	91 (51%)	64 (36%)
3 – 3.75 Freshmen	44 (20%)	15 (7%)	23 (10%)	43 (19%)
3 – 3.75 Seniors	76 (42%)	43 (24%)	56 (31%)	79 (44%)
4 Freshmen	6 (3%)	1 (0%)	0 (0%)	3 (1%)
4 Seniors	22 (12%)	17 (9%)	12 (7%)	8 (4%)
Grand Total Freshmen	223 (100%)	223 (100%)	223 (100%)	223 (100%)
Grand Total Seniors	179 (100%)	179 (100%)	179 (100%)	179 (100%)

Freshman Baseline/Senior Exiting Comparisons
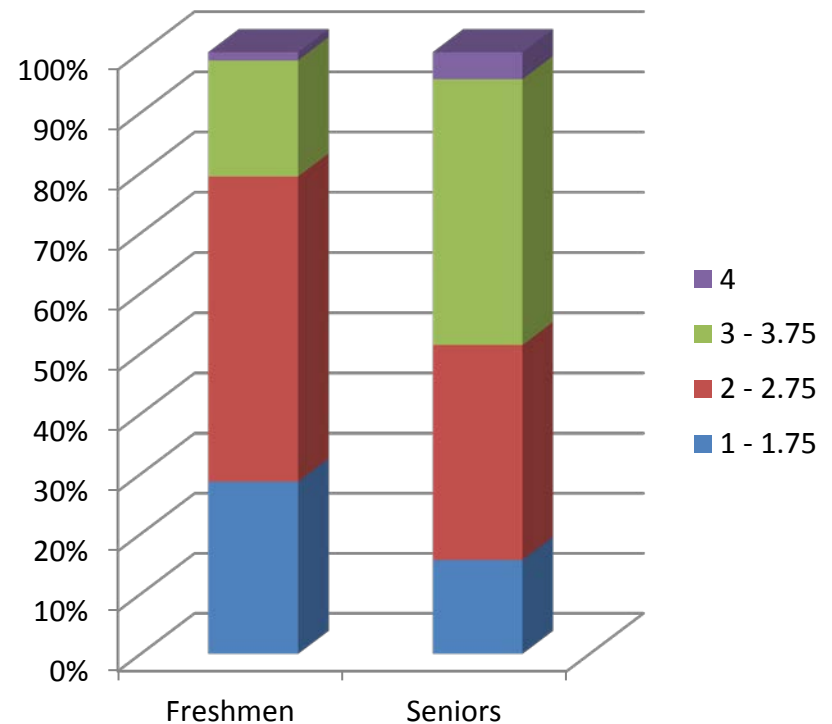## Freshman *n* = 223; Senior *n* = 179
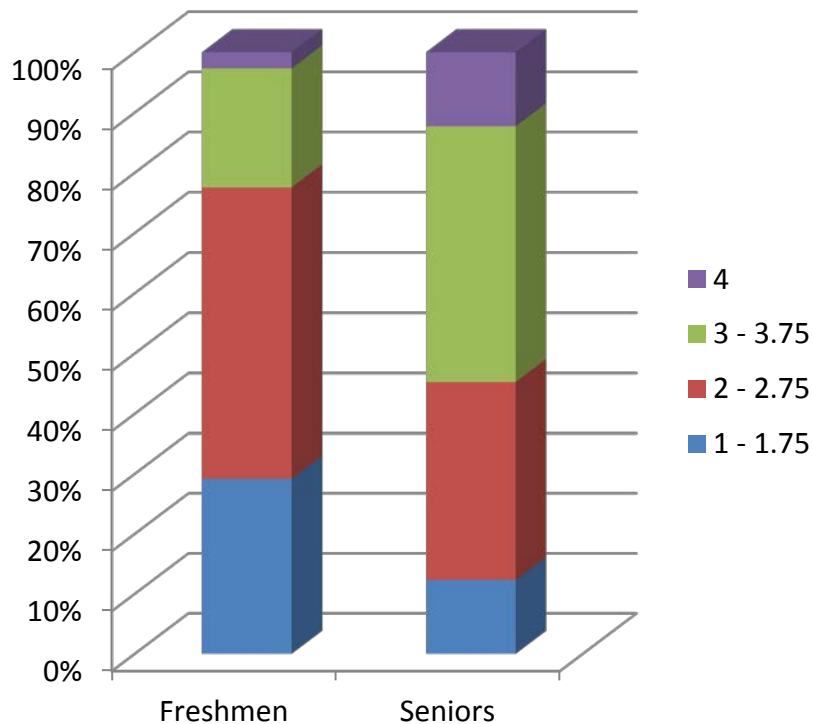
Freshman Baseline/Senior Exiting Comparisons
Freshman *n* = 223; Senior *n* = 179

Diction
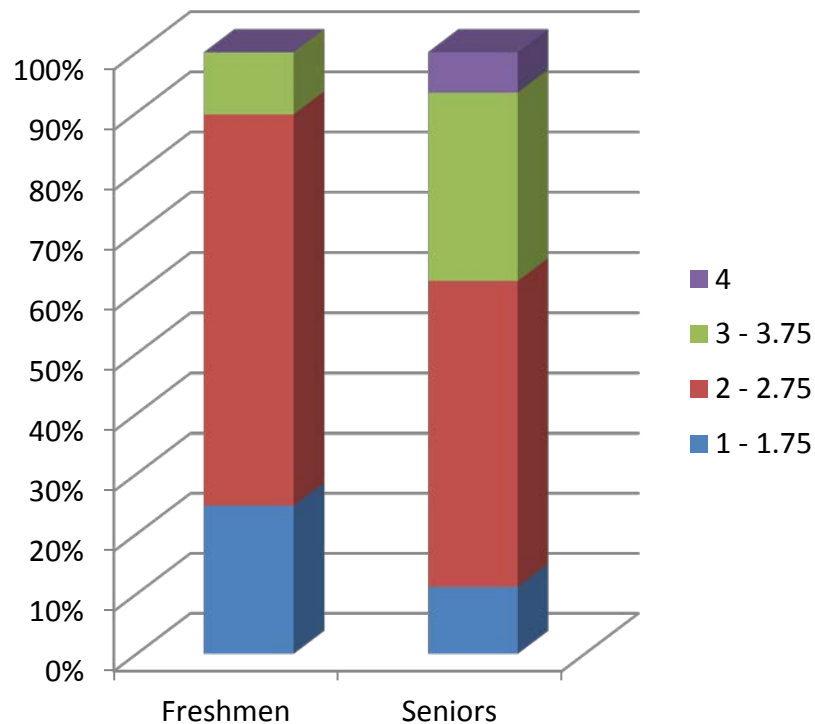
Syntax/Mechanics

Freshman Baseline/Senior Exiting Assessment Inter-Rater Agreement Results

Trait/ Agreement	Organization (Kappa = .161)	Genre (Kappa = .290)	Diction (Kappa = .210)	Syntax/Mechanics (Kappa = .215)
Agree	160 (40%)	202 (50%)	200 (50%)	184 (46%)
Difference = 1 point or less	195 (49%)	165 (41%)	174 (43%)	178 (44%)
Difference = 1.5 to 2 points	43 (11%)	31 (8%)	28 (7%)	38 (9%)
Difference = 2.5 to 3 points	4 (1%)	4 (1%)	0 (0%)	2 (0%)
Total	402 (100%)	402 (100%)	402 (100%)	402 (100%)

Comparison of Freshman Baseline and First-Year Seminar (FYS) Assessments

Academic Year 2013 - 2014

Review Procedures

- The review team had FYS final assessments for 151 of the 223 freshmen who had completed baseline assessments during Week of Welcome. FYS assessments were evaluated across the same nine criteria (traits) used to score freshman baseline assessments. Scoring methodology also was the same.

Freshman Baseline/FYS Comparisons

Mean Scores on a scale of 1 – 4, with 4 being the highest possible score
n = 151 (Information Needed *ns*)
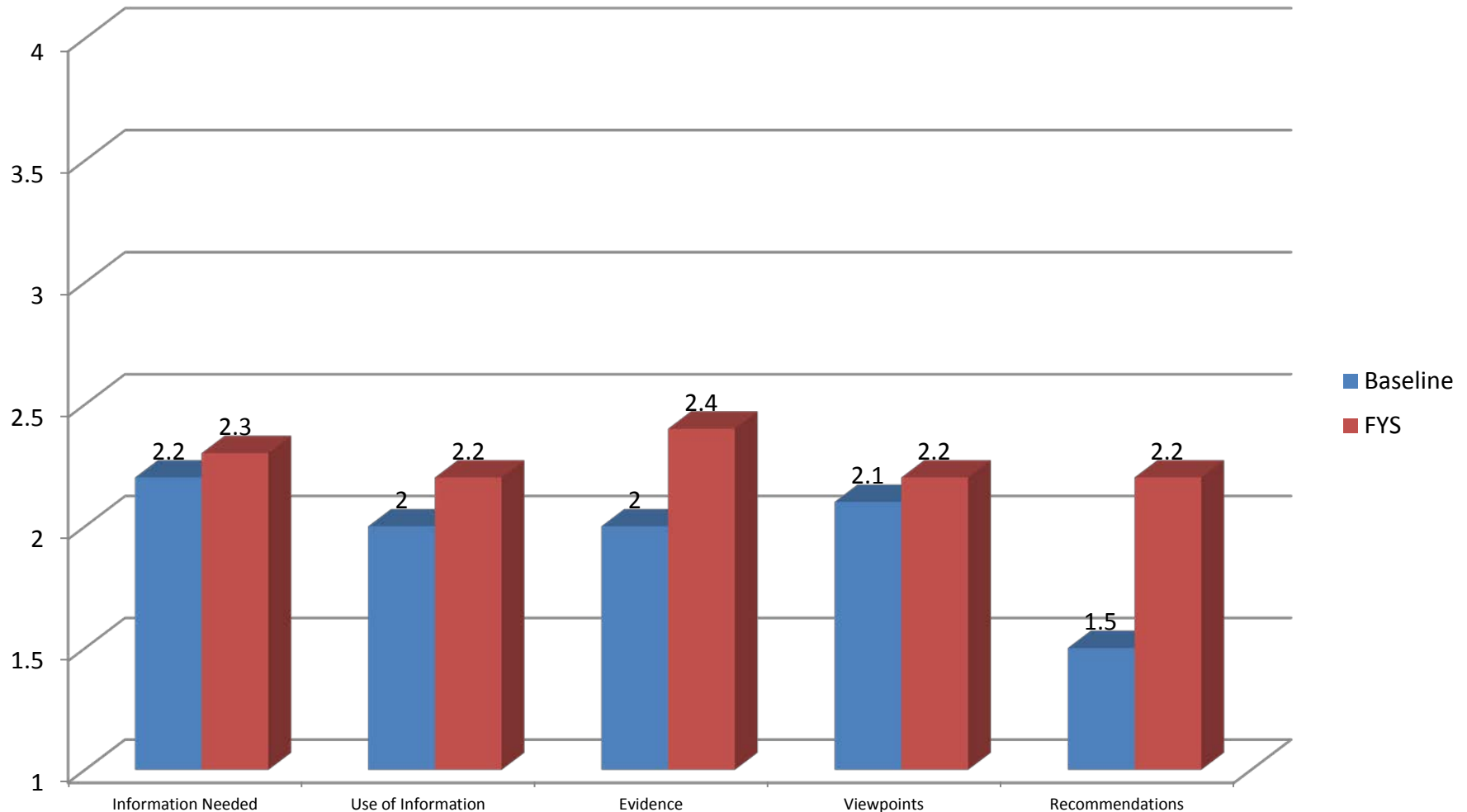
Freshman Baseline/FYS Comparisons
n = 151

| Trait/ Performance Level	Info Needed	Use of Info	Evidence	Viewpoints	Recommendations
1 – 1.75 Baseline	23 (15%)	68 (45%)	46 (30%)	47 (31%)	111 (74%)
1 – 1.75 FYS	17 (11%)	47 (31%)	20 (13%)	23 (15%)	54 (36%)
2 – 2.75 Baseline	99 (66%)	56 (37%)	89 (59%)	76 (50%)	38 (25%)
2 – 2.75 FYS	96 (64%)	56 (37%)	86 (57%)	105 (70%)	72 (48%)
3 – 3.75 Baseline	28 (19%)	24 (16%)	16 (11%)	27 (18%)	2 (1%)
3 – 3.75 FYS	37 (25%)	45 (30%)	42 (28%)	22 (15%)	25 (17%)
4 Baseline	1 (1%)	3 (2%)	0 (0%)	1 (1%)	0 (0%)
4 FYS	1 (1%)	3 (2%)	3 (2%)	1 (1%)	0 (0%)
Grand Total Baseline	151 (100%)	151 (100%)	151 (100%)	151 (100%)	151 (100%)
Grand Total FYS	151 (100%)	151 (100%)	151 (100%)	151 (100%)	151 (100%)
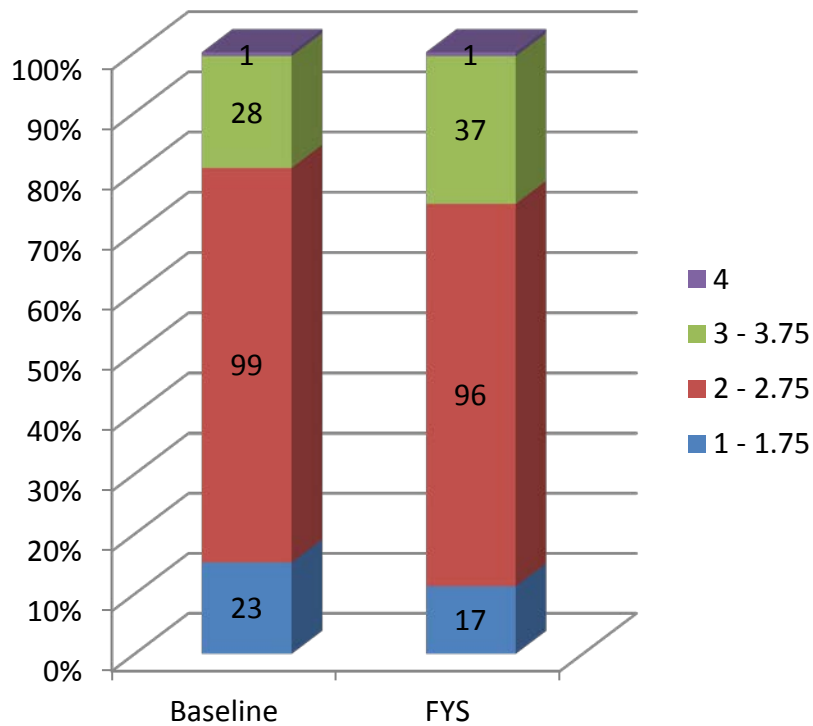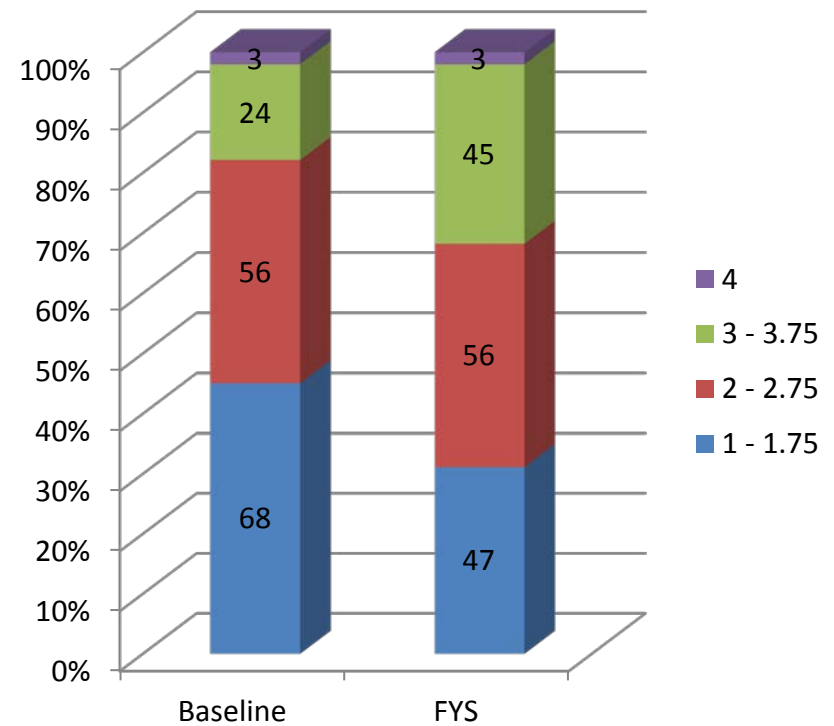
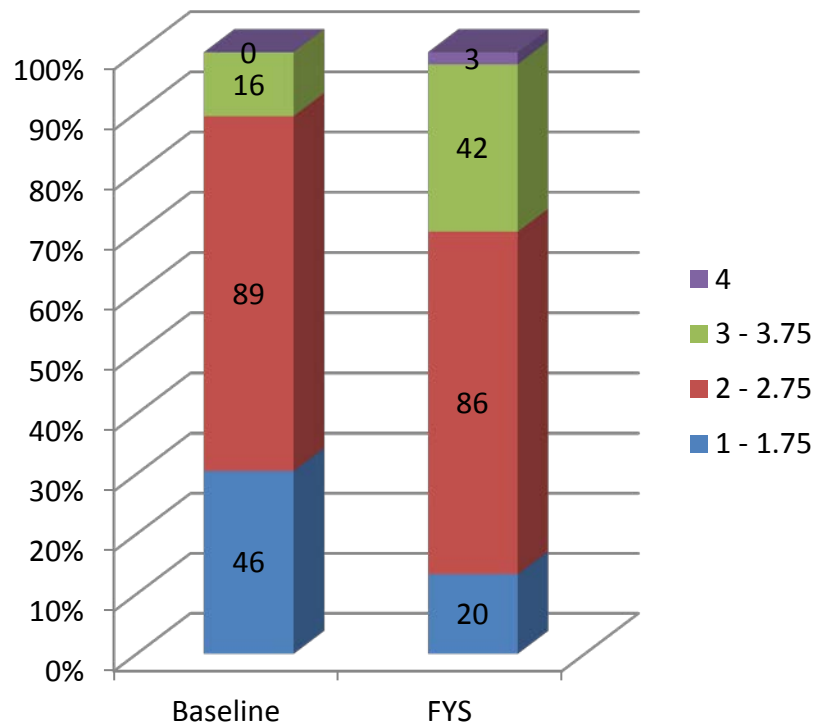# Freshman Baseline/FYS Comparisons
n = 151

Freshman Baseline/FYS Comparisons
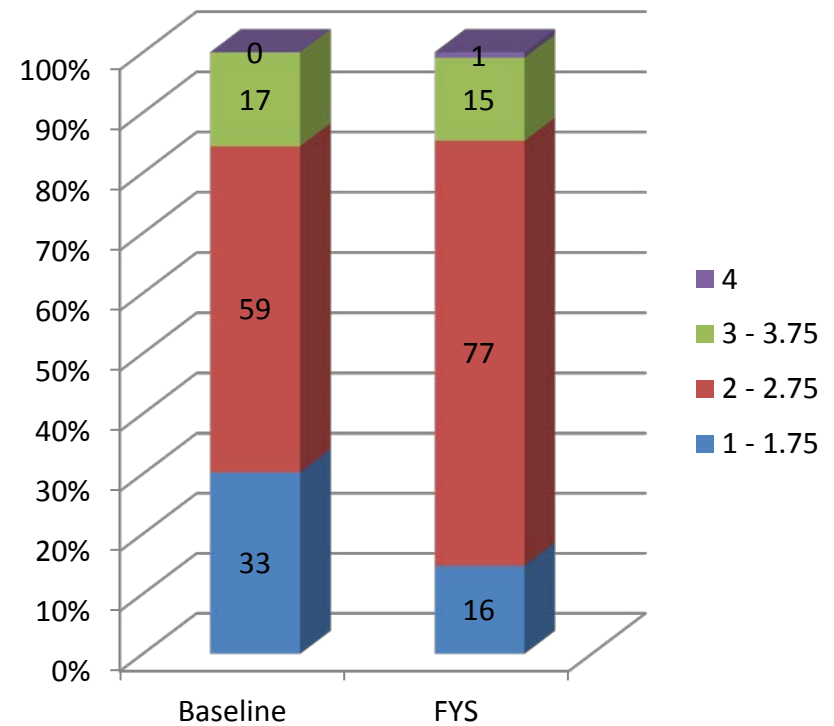
n = 151

Freshman Baseline/FYS Comparisons
n = 151

Recommendations

Chart legend:
- 4
- 3 - 3.75
- 2 - 2.75
- 1 - 1.75

Baseline: 111, 38, 2
FYS: 54, 72, 25
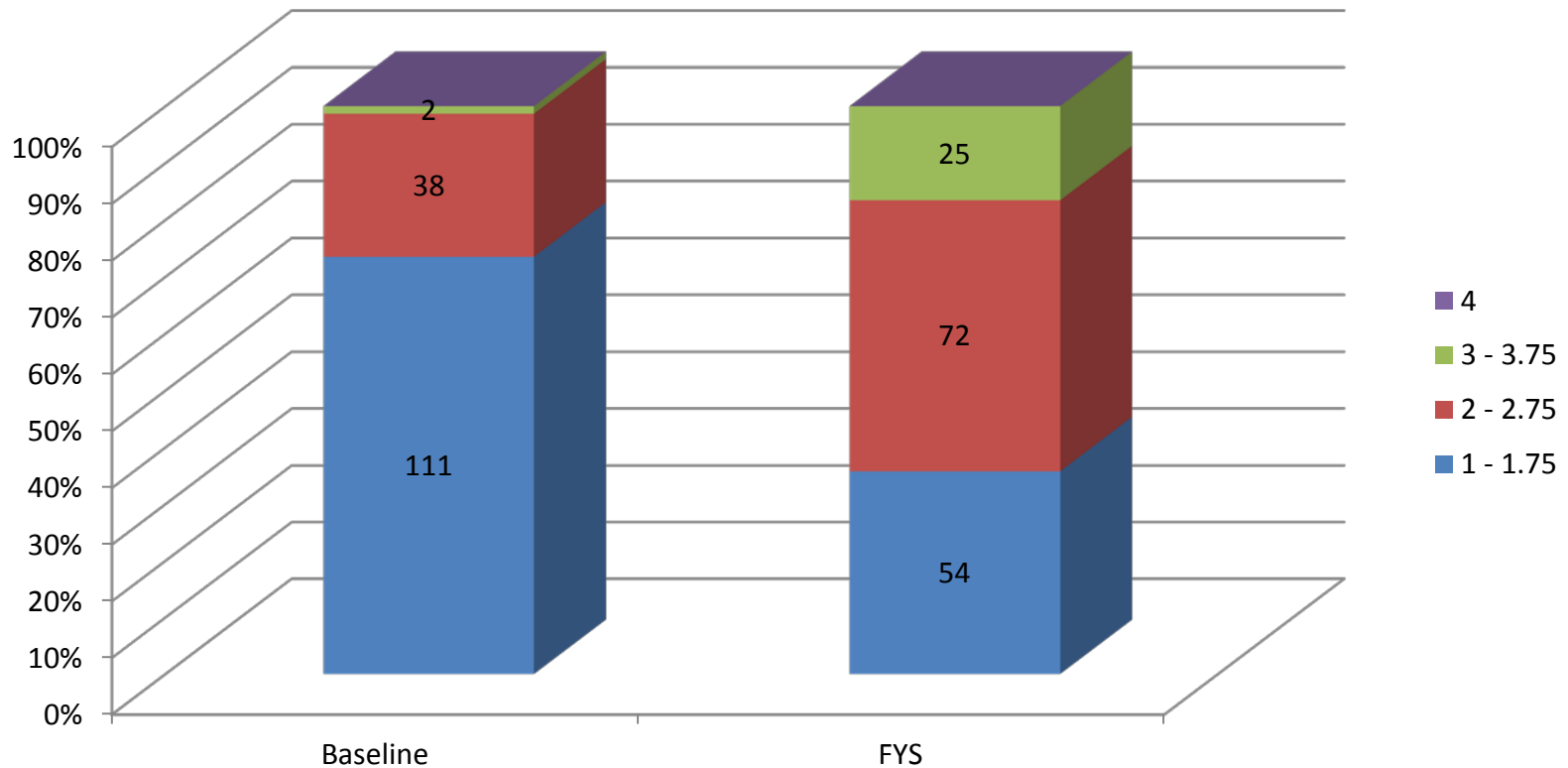
FYS Inter-Rater Agreement Results

Trait/ Agreement	Info Needed (Kappa = .385)	Use of Info (Kappa = .428)	Evidence (Kappa = .200)	Viewpoints (Kappa = .262)	Recommendations (Kappa = .047)
Agree	96 (64%)	91 (60%)	74 (49%)	89 (59%)	56 (37%)
Difference = 1 point or less	52 (34%)	55 (36%)	71 (47%)	55 (36%)	81 (54%)
Difference = 1.5 to 2 points	2 (1%)	5 (3%)	5 (3%)	6 (4%)	13 (9%)
Difference = 2.5 to 3 points	1 (1%)	0 (0%)	1 (1%)	1 (1%)	1 (1%)
Total	151 (100%)	151 (100%)	151 (100%)	151 (100%)	151 (100%)

Freshman Baseline/FYS Comparisons

Mean Scores on a scale of 1 – 4, with 4 being the highest possible score
n = 151 (Syntax/Mechanics *ns*)
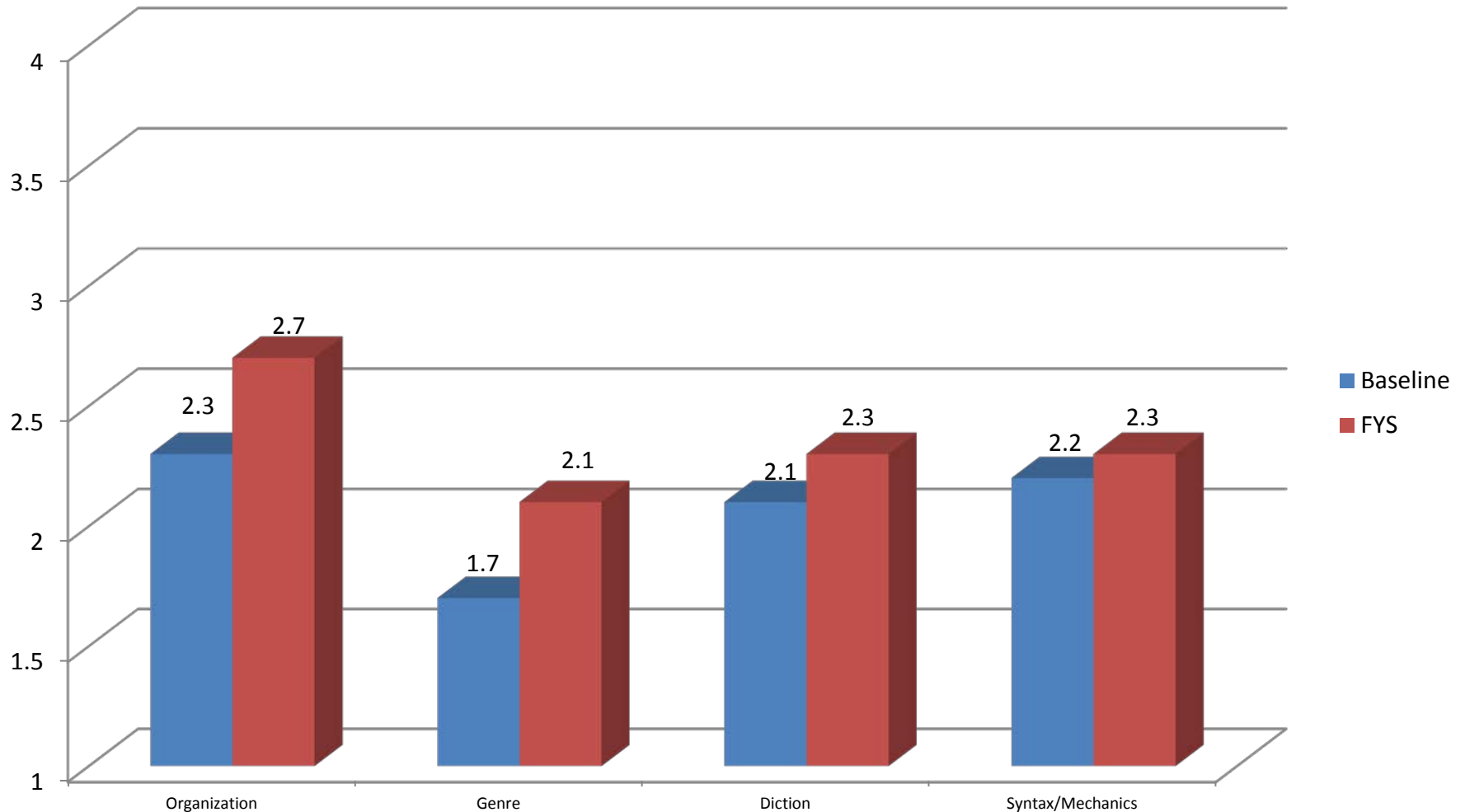
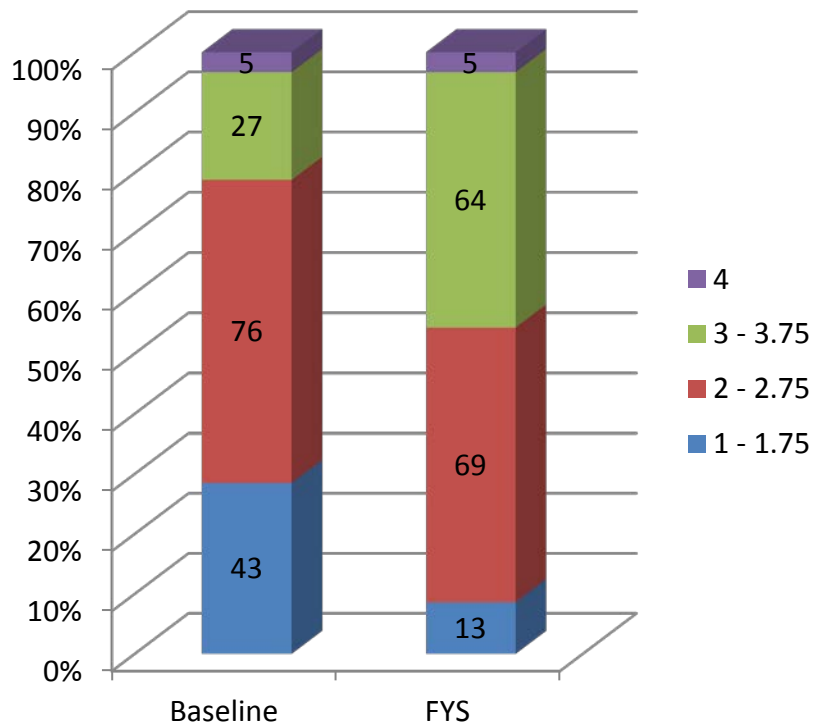# Freshman Baseline/FYS Comparisons
n = 151

| Trait/ Performance Level	Organization	Genre	Diction	Syntax/Mechanics
1 – 1.75 Baseline	43 (28%)	89 (59%)	33 (22%)	39 (26%)
1 – 1.75 FYS	13 (9%)	74 (49%)	18 (12%)	33 (22%)
2 – 2.75 Baseline	76 (50%)	49 (32%)	101 (67%)	77 (51%)
2 – 2.75 FYS	69 (46%)	36 (24%)	102 (68%)	78 (52%)
3 – 3.75 Baseline	27 (18%)	12 (8%)	17 (11%)	33 (22%)
3 – 3.75 FYS	64 (42%)	34 (23%)	30 (20%)	39 (26%)
4 Baseline	5 (3%)	1 (1%)	0 (0%)	2 (1%)
4 FYS	5 (3%)	7 (5%)	1 (1%)	1 (1%)
Grand Total Baseline	151 (100%)	151 (100%)	151 (100%)	151 (100%)
Grand Total FYS	151 (100%)	151 (100%)	151 (100%)	151 (100%)
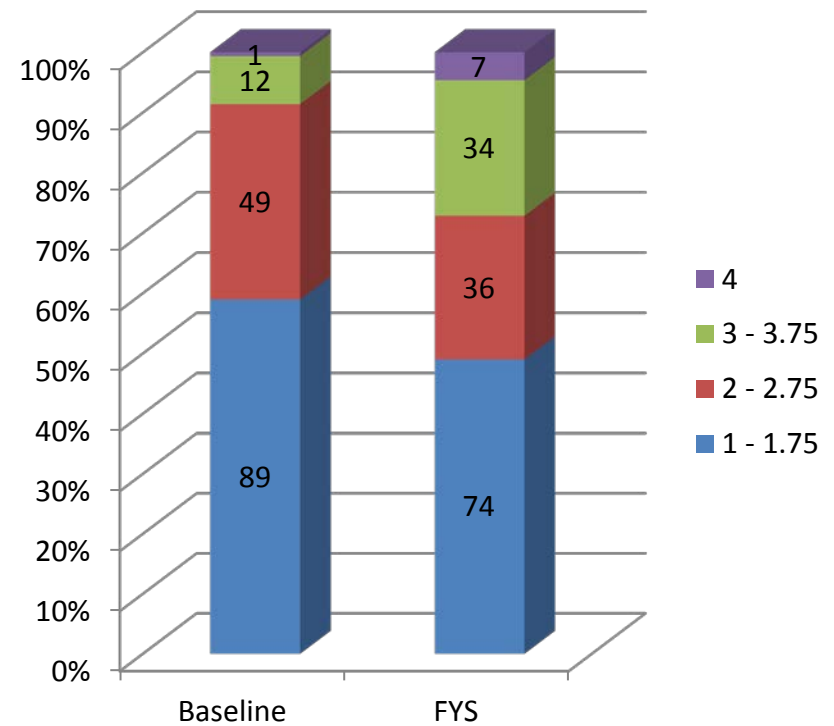
Freshman Baseline/FYS Comparisons
n = 151
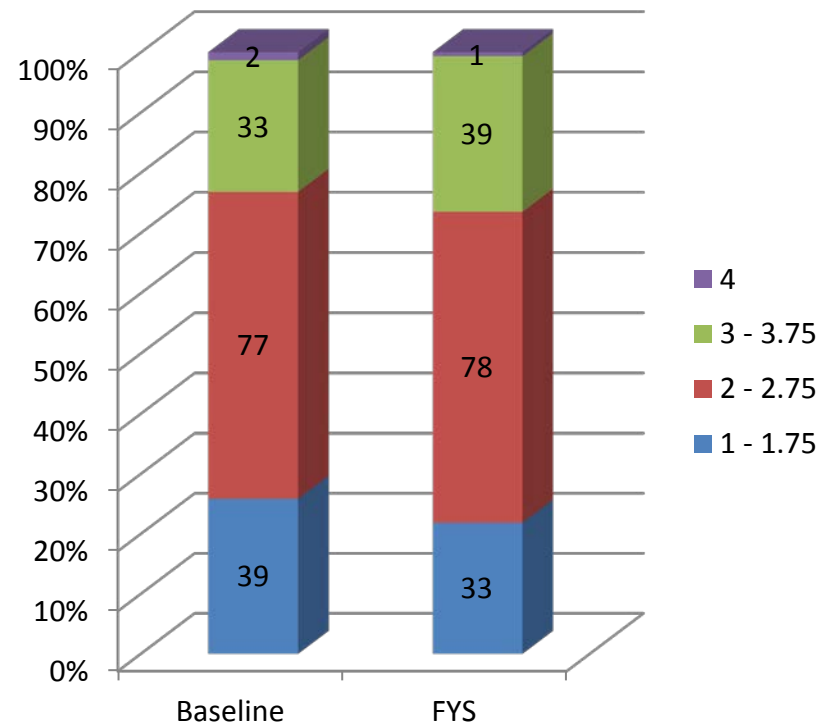
Freshman Baseline/FYS Comparisons
n = 151
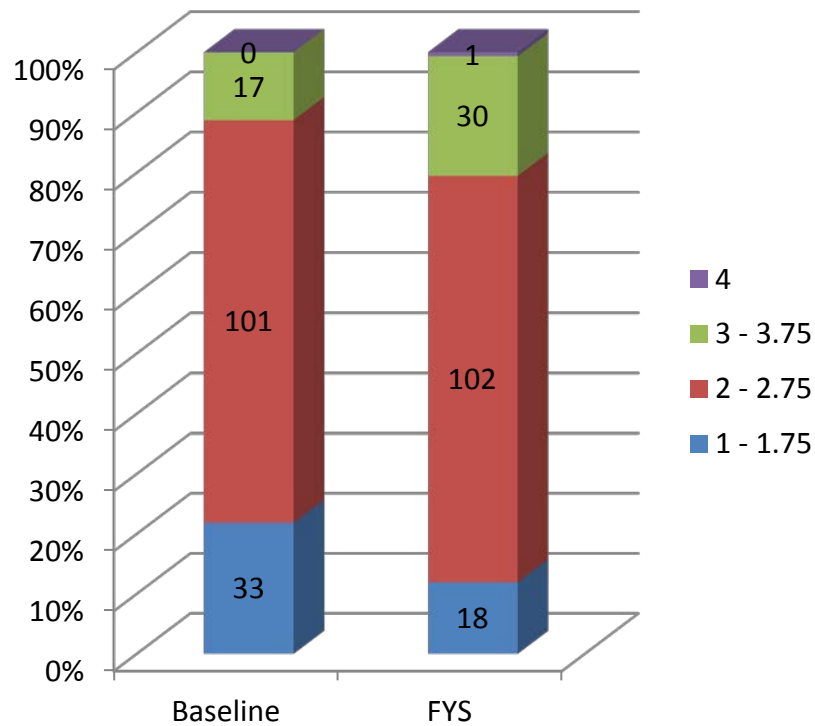
FYS Inter-Rater Agreement Results

Trait/ Agreement	Organization (Kappa = .213)	Genre (Kappa = .349)	Diction (Kappa = .183)	Syntax/Mechanics (Kappa = .177)
Agree	76 (50%)	81 (54%)	81 (54%)	69 (46%)
Difference = 1 point or less	67 (44%)	65 (43%)	65 (43%)	57 (38%)
Difference = 1.5 to 2 points	7 (5%)	5 (3%)	5 (3%)	25 (17%)
Difference = 2.5 to 3 points	2 (1%)	0 (0%)	0 (0%)	0 (0%)
Total	151 (100%)	151 (100%)	151 (100%)	151 (100%)