Comparison of Freshman Baseline with First Year Seminar Assessment Results

Academic Year 2022 – 2023

Summer Assessment Team Members: Marie Archambault, Clinton Brown, Kim DeTardo-Bora, Victor Fet, Marty Laubach, Leslie Dawn Quick, and Anita Walz

Summer Assessment Support Staff: Mary Beth Reynolds, Adam Russell, Mary Welch, and Diana Adams

Executive Summary

Background

Recommendations from the 2022 Assessment Team (current status is in red)

The Summer Assessment Team made the following recommendations:

1. That evidence documents attached to FYS scenarios be evaluated for equivalence of type, length, and complexity across scenarios. We noted in our discussions in May 2023 that the reason for this recommendation – that FYS students had to complete these assessments, including reading the evidentiary documents, within a two-hour window – no longer exists. They now have a week to complete these assessments.
2. That creators of baseline and FYS scenarios consider validated scales that students could use when assessing documents for creditability and relevance. This recommendation was not implemented.
3. That students be asked to provide a two-sentence summary regarding why they have judged the credibility and relevance of each document as they have. Please refer to recommendations from this year's Summer Assessment Team at the end of this document.
4. That the Summer Assessment Team review the current rubric before starting the assessment in summer 2023. This evaluation was completed and resulted in some changes to the rubric, as noted in the report.
5. That we include a more comprehensive evaluation of information literacy in our ratings, e.g., if students say they're using peer-reviewed journals, note where that would that fall on the rubric scale. This recommendation was not implemented; however, the Summer Assessment Team emphasized that, depending on the critical thinking scenario used, evidence other than that from peer-reviewed studies could be deemed credible and relevant.
6. That FYS instructors consider having students use the same convention/format to make their recommendations. This would place all students on the same playing field for achievement on *Communication Fluency*: convention/format; however, it is worth noting that *Communication Fluency* is not one of the outcomes of FYS, and we realize that FYS instructors might have a

pedagogical reason for using different convention/formats for FYS scenarios. We note again the Communication Fluency is not an outcome of FYS.

7. That we consider using the same rubrics for baseline, FYS, and senior capstone projects. We discussed this recommendation again after scoring the senior capstone assessments in 2023. While it was recommended, we are considering shifting the senior assessment to programmatic senior assignments that align with one or more of the Baccalaureate Degree Profile outcomes. This would entail using those rubrics for assessment.
8. As was recommended last year, the Office of Assessment and Quality Initiatives should continue to provide and distribute shorter reports in more digestible formats. We recommend that these reports be disseminated campus-wide through the Assessment Newsletter. This will be a priority for academic year 2023-2024.

Procedures for the 2023 Assessment

General Procedures

In August 2022, 1,271 incoming freshmen at Marshall University at least appeared to have uploaded baseline assessments into Blackboard as part of their assignments for Freshman First Class (UNI 100). These assessments required students 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 (Critical) Thinking*, and *Communication Fluency*. As part of Marshall's mandatory First Year Seminar in Critical Thinking (FYS), students completed assessments that mirrored those they finished as incoming freshmen, with 971 FYS assessments uploaded into Blackboard. To obtain a sample of matched pairs of baseline and FYS assessments, we began by examining a random sample of 362 artifacts, each of which had a match from both baseline and FYS queues. We then examined each to ensure that they had the appropriate artifact uploaded in both their baseline and FYS queues. This process yielded a total of 175 matched baseline/FYS pairs. Please note that our sample represented 14% of uploaded baseline and 18% of uploaded FYS assessments. During the Assessment Team's review, we discovered that one baseline artifact and its matching FYS artifact were in fact completed and submitted by different students who shared the same name. We note that Blackboard is set up in a way that does not allow positive identification of students until the final score download, thus hiding this information from both the administrator and the reviewers until that time. An additional student in our sample, who completed a baseline assessment, uploaded the FYS assessment instructions rather than the work produced required. This reduced the usable number of matched pairs to 173.

In May 2023, a group of seven faculty representing several academic colleges from across the university evaluated the baseline/FYS sample using a rubric that allowed them to score each artifact across eight criteria (traits). These traits included information needed and source acknowledgment (*Information Literacy*), evidence, viewpoints, and recommendation/position (I*nquiry-Based [Critical] Thinking*), and development, convention/format, and communication style (*Communication Fluency*). This project was coordinated by the Office of Assessment and Quality Initiatives.

Each assessment had two independent raters. Please see the supporting documentation that follows this summary for a detailed explanation of scoring procedures.

Results and Analysis

Comparison of Freshman Baseline to Results at the End of FYS

The baseline and FYS means (and standard deviations) for the students in the sample with scorable baseline and FYS exams are reported below. Please note that, for students with scorable baseline and FYS (i.e., pre-post) assessments, *paired-samples t-tests* using adjusted alpha levels to control for Type I error (.025 for I*nformation literacy*), (.017 for *Inquiry-Based [Critical] Thinking*), and (.017 for *Communication Fluency*) showed significant mean differences between freshman baseline and FYS results for one trait (source acknowledgment) of *Information Literacy* and for one trait (development) of *Communication Fluency*. Overall, the students in this sample did not improve significantly between baseline and FYS in any traits of Inquiry-Based (Critical) Thinking. We further note that *Communication Fluency* is not an outcome of FYS.

Outcome	Trait	Baseline Mean (*SD*)	FYS Mean (*SD*)	Statistical Significance
Information Literacy	**Information Needed**	2.199 (0.5412)	2.260 (0.6506)	*t(172)* = -1.177, *p* = .241
	Source Acknowledgment	2.043 (0.8245)	2.251 (0.7825)	*t(172)* = -2.675, *p* = .008
Inquiry-Based (Critical) Thinking	**Evidence**	2.208 (0.7353)	2.257 (0.7145)	*t(172)* = -0.850, *p* = .396
	Viewpoints	1.902 (0.4872)	1.948 (0.5500)	*t(172)* = -0.953, *p* = .342
	Recommendation/ Position	2.230 (0.6922)	2.304 (0.6683)	*t(172)* = -0.095, *p* = .924
Communication Fluency	**Development**	2.208 (0.7334)	2.361 (0.6455)	*t(172)* = -2.730, *p* = .007
	Convention/Format	2.535 (0.7165)	2.587 (0.7594)	*t(172)* = -0.716, *p* = .475
	Communication Style	2.682 (0.5552)	2.723 (0.5293)	*t(172)* = -0.851, *p* = .396

A frequency analysis also showed the following increases in students scoring between 2.5 and 4.0 on the rubric between baseline and FYS. Please see the supporting documentation following this summary for additional information.

Outcome	Trait	Percentage Gain in Students Scoring 2.5 to 4.0 from Baseline to FYS
Information Literacy	**Information Needed**	9%
	Source Acknowledgment	13%
Inquiry-Based (Critical) Thinking	**Evidence**	8%
	Viewpoints	8%
	Recommendation/Position	10%
Communication Fluency	**Development**	11%
	Convention/Format	7%
	Communication Style	2%

This year's results showed a significant difference in performance based on scenario used for the FYS assessments for one trait (recommendation/position) of *Inquiry-Based [Critical] Thinking*, and for two traits (convention/format, and communication style) of *Communication Fluency*. On convention/format, students scored significantly lower on GMO Foods than on the other two scenarios (Online Gaming and Social Media). On communication style, students scored significantly higher on Online Gaming than on either GMO Foods or Social Media. There was not a significant difference in mean scores between Online Gaming and Social Media. On recommendation/position mean scores were significantly lower for GMO Foods than for Online Gaming and Social Media.

Gain scores between students in our sample who completed FYS in fall 2022 (n = 56) and those who completed FYS in spring 2023 (n = 117) differed significantly on only one outcome trait, *Communication Fluency* (convention/format), with students enrolled in the fall (mean gain = +0.438) outperforming students enrolled in the spring (mean gain = -0.133), t (171) = 3.813; p < .001. Again, we note that *Communication Fluency* is not an outcome of FYS. Please refer to the supporting documentation for additional detail.

Conclusions

Although we have not performed statistical analyses to compare the results across years, we note the following data patterns using this report (academic year 2022-2023) as a reference point.

Baseline Mean Results (Communication Fluency [CF] is not an outcome of FYS)

Outcome: Rubric Trait	Fall 2022 (Reference)	Fall 2021	Fall 2020	Fall 2019	Fall 2018	Fall 2017	Fall 2016
IL: Information Needed	**2.199**	2.108 **Lower**	2.377 Higher	2.116 **Lower**	2.34 Higher	2.32 Higher	2.37 Higher
IL: Source Acknowledgment	**2.043**	2.029 **Lower**	2.241 Higher	1.355 **Lower**	1.77 **Lower**	1.94 **Lower**	2.12 Higher
CT: Evidence	**2.208**	2.141 **Lower**	2.322 Higher	1.798 **Lower**	2.02 **Lower**	2.23 Higher	2.19 **Lower**
CT: Viewpoints	**1.902**	1.962 Higher	2.048 Higher	1.847 **Lower**	1.87 **Lower**	2.09 Higher	2.27 Higher
CT: Recommendation/ Position	**2.230**	2.317 Higher	2.462 Higher	2.252 Higher	2.09 **Lower**	2.19 **Lower**	1.96 **Lower**
CF: Development	**2.208**	2.199 Lower	2.317 Higher	2.029 Lower	2.02 Lower	2.12 Lower	2.06 Lower
CF: Convention/ Format	**2.535**	2.407 Lower	2.513 Lower	2.306 Lower	1.95 Lower	1.59 Lower	1.85 Lower
CF: Communication Style	**2.682**	2.587 Lower	2.663 Lower	2.393 Lower	2.37 Lower	2.36 Lower	2.35 Lower

FYS Mean Results (Communication Fluency [CF] is not an outcome of FYS)

Outcome: Rubric Trait	2022-2023 (Reference)	2021-2022	2020-2021	2019-2020	2018-2019	2017-2018	2016-2017
IL: Information Needed	**2.260**	2.351 **Higher**	2.525 **Higher**	2.519 **Higher**	2.52 **Higher**	2.48 **Higher**	2.39 **Higher**
IL: Source Acknowledgment	**2.251**	2.279 **Higher**	2.457 **Higher**	2.471 **Higher**	2.27 **Higher**	2.45 **Higher**	2.37 **Higher**

Outcome: Rubric Trait	2022-2023 (Reference)	2021-2022	2020-2021	2019-2020	2018-2019	2017-2018	2016-2017
CT: Evidence	**2.257**	2.365 **Higher**	2.503 **Higher**	2.360 **Higher**	2.25 Same	2.36 **Higher**	2.32 **Higher**
CT: Viewpoints	**1.948**	2.112 **Higher**	2.106 **Higher**	2.198 **Higher**	2.01 **Higher**	2.25 **Higher**	2.16 **Higher**
CT: Recommendation/ Position	**2.304**	2.519 **Higher**	2.550 **Higher**	2.471 **Higher**	2.25 Lower	2.41 **Higher**	2.36 **Higher**
CF: Development	**2.361**	2.452 Higher	2.513 Higher	2.446 Higher	2.24 Lower	2.25 Lower	2.28 Lower
CF: Convention/ Format	**2.587**	2.683 Higher	2.774 Higher	2.566 Lower	2.28 Lower	1.89 Lower	2.15 Lower
CF: Communication Style	**2.723**	2.728 Higher	2.663 Lower	2.591 Lower	2.50 Lower	2.38 Lower	2.44 Lower

FYS Mean Results (Note: Significant = FYS significantly higher than baseline; NS = no significant difference between baseline and FYS)

Outcome: Rubric Trait	2022-2023 (Reference)	2022-2023 (Reference)	2021-2022	2020-2021	2019-2020	2018-2019	2017-2018	2016-2017
IL: Information Needed	**2.260**	**NS**	**Significant**	**Significant**	**Significant**	**Significant**	**Significant**	**NS**
IL: Source Acknowledgment	**2.251**	**Significant**	**Significant**	**Significant**	**Significant**	**Significant**	**Significant**	**Significant**
CT: Evidence	**2.257**	**NS**	**Significant**	**Significant**	**Significant**	**Significant**	**NS**	**NS**
CT: Viewpoints	**1.948**	**NS**	**Significant**	**NS**	**Significant**	**Significant**	**Significant**	**NS**
CT: Recommendation/ Position	**2.304**	**NS**	**Significant**	**NS**	**Significant**	**Significant**	**Significant**	**Significant**

Using the fall 2022 baseline numbers as a reference point, we conclude that this year's results are not due to higher than usual baseline scores, but rather to lower FYS scores than in past years. Although we did not include data from 2013, 2014, or 2015, we note that significant improvement between baseline and FYS was seen for at least one trait of *critical thinking* in those years as well. Given these data, we are concerned that this was the first year since 2013 that students did not improve significantly in their performance on any trait of critical thinking.

Recommendations from the 2023 Assessment Team

The Summer Assessment Team made the following recommendations:

1. That we reflect on the original purpose of the course we call "FYS," whose name is "First Year Seminar in **Critical Thinking.** We were concerned that this is the first year since we have been assessing change in outcomes related to *Information Literacy* and *Critical Thinking* that we saw no significant difference between student performance on their baseline assessments and assessments at the conclusion of their FYS experience. We recommend that additional support be provided to instructors to help them craft their pedagogy to focus on critical thinking during this course. This should be done by returning a faculty member to the position of FYS coordinator.
2. That students be asked to provide a two-sentence summary regarding why they have judged the credibility and relevance of each document as they have. This recommendation is repeated from last year.

Supporting Documentation

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

Academic Year 2022 - 2023

Review Procedures

- One hundred seventy-five (175) FYS critical thinking artifacts were matched with 175 baseline critical thinking artifacts. This number represented 18% of the 971 FYS artifacts and 14% of the 1,291 baseline artifacts uploaded to Blackboard.

- During the evaluation we discovered that one baseline artifact and its matching FYS artifact were in fact rendered by different students who shared the same name. An additional student in our sample who completed a baseline assessment uploaded FYS assessment instructions rather than the work product required. This reduced the usable matched pairs for comparison of means and frequencies to 173.

Review Procedures Continued

- Each assessment had two independent raters and scores were determined in the following manner:
  - If raters assigned the same score, that became the score for the artifact.
  - If raters' scores differed by one point, 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.
  - 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.
  - 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. (For this review, all raters were able to reconcile disagreements, so third raters were not needed).

Interrater Reliability

- We conducted interrater reliability analyses using the Cohen's Kappa statistical procedure. In so doing, we used the following rules, similar to those suggested by Stellmack, Kohneim-Kalkstein, Manor, Massey, & Schmitz (2009):
  - Since our scoring procedure was to average final scores between two raters when scores differed by only one point, we used that averaged score (e.g., 1.5) as the score for both raters, counting it as an agreement in the interrater reliability analysis.
  - For scores that were two or more points apart, the original score of each reviewer was used in the analysis. Therefore, these scores were counted as disagreements.

Rubric Used for Scoring

Baseline/FYS Assessment Rubric – Summer 2023 – updated 5-8-2023

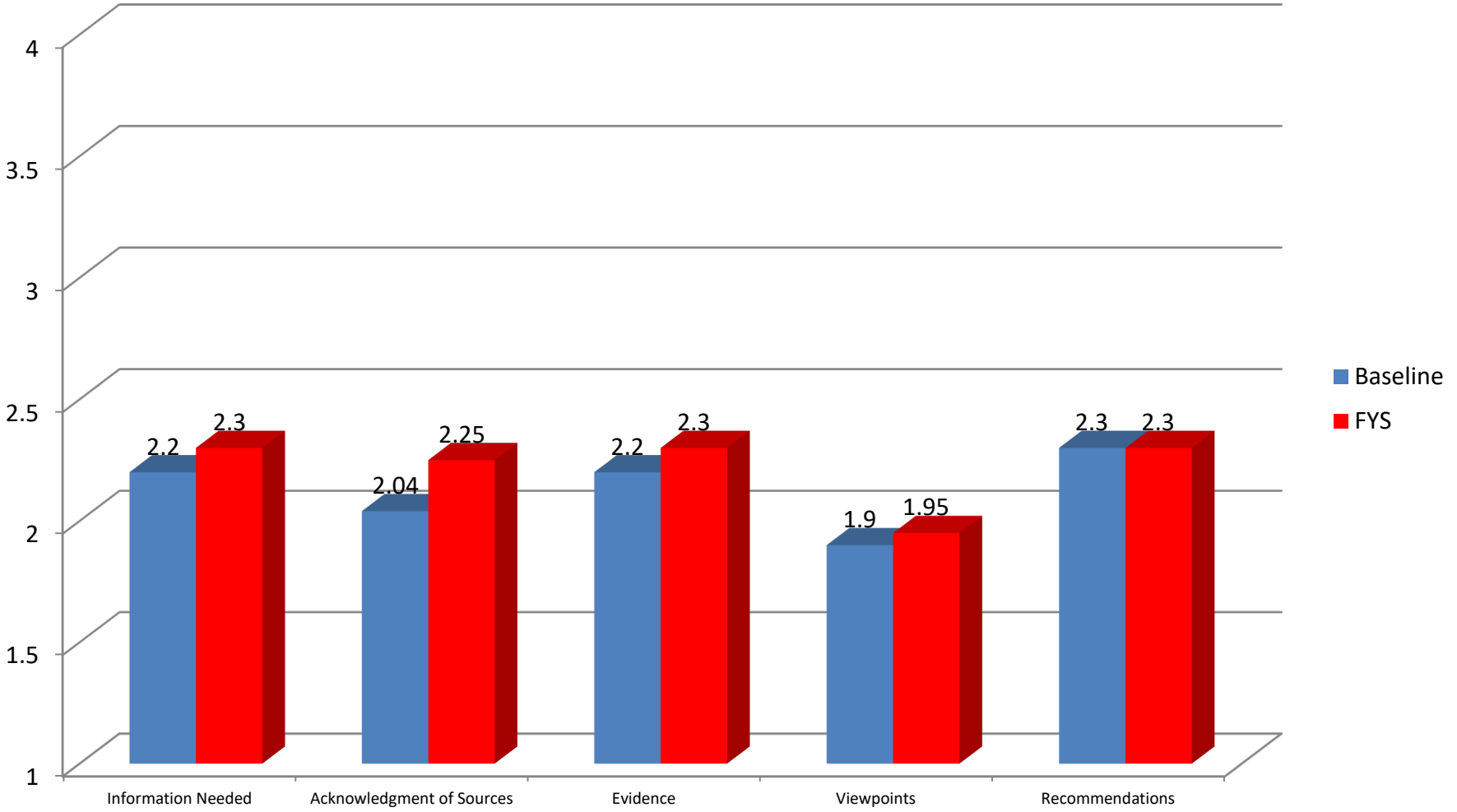
Outcomes	Traits	Performance Levels			
		1	2	3	4
Information Literacy	**Information Needed**	Does not acknowledge or assess the need for more information.	Acknowledges the need for more information but does not identify research methods/sources (or those identified are not feasible) that would address unanswered questions.	Assesses the need for more information and recommends general research methods/sources (that are feasible) that would address some unanswered questions.	Assesses the need for more information and recommends specific research methods/sources (that are feasible) that would address most unanswered questions.
	Source Acknowledgment	Fails to acknowledge sources from the DL.	Indirectly/vaguely acknowledges sources of information from the DL.	Clearly acknowledges relevant sources of information from the DL.	Integrates relevant information from the DL. Acknowledges sources used.
Inquiry-Based Thinking	**Evidence**	Disregards or misunderstands evidence from the DL.	Insufficient evidence is taken from sources (e.g., only one or two pieces of evidence) in the DL or evidence is used without appropriate interpretation/evaluation (i.e., poor job).	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 (i.e., adequate job).	Evidence is taken from relevant and valid sources in the DL with enough interpretation/evaluation to develop a coherent analysis or synthesis (i.e., good/excellent job).
	Viewpoints	Ignores 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 (take a position) or makes a recommendation (takes a position), but does not justify it in any way.	Recommendation/position is justified, but does not acknowledge different sides of the issue.	Recommendation/position is justified and takes into account different sides/complexities of the issue.	Recommendation/position 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 their ideas.	Shows some development of ideas.	Shows a strong, but perhaps somewhat incomplete, development of ideas.	Produces a document in which the ideas have been fully developed.
	Convention/Format	Demonstrates minimal attention to basic organization and presentation and stylistic conventions.	Demonstrates some awareness of basic organization, content, and presentation and stylistic conventions.	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 impedes meaning because of errors in usage/mechanics.	Uses language that generally conveys meaning to readers, although errors in usage/mechanics impedes smooth reading of the document.	Uses straightforward language that conveys meaning to readers. The language in the document has few errors.	Uses language that skillfully communicates meaning to readers with clarity and fluency, and is virtually error-free.

Freshman Baseline/FYS Comparisons

Mean Scores on a scale of 1 – 4, with 4 being the highest possible score

n = 173

Mean differences were statistically significant for *Acknowledgement of Sources;* mean differences for all other traits below were not statistically significant.



Freshman Baseline/FYS Comparisons

n = 173

Trait/ Performance Level	Info Needed (Baseline)	Info Needed (FYS)	Acknowledgment of Sources (Baseline)	Acknowledgment of Sources (FYS)
1.0	10 (6%)	14 (8%)	40 (23%)	32 (18%)
1.5 – 2.0	96 (55%)	79 (46%)	61 (35%)	46 (27%)
2.5 – 3.0	61 (35%)	72 (42%)	60 (35%)	88 (51%)
3.5 – 4.0	6 (3%)	8 (5%)	12 (7%)	7 (4%)
Totals	**173**	**173**	**173**	**173**

Freshman Baseline/FYS Comparisons

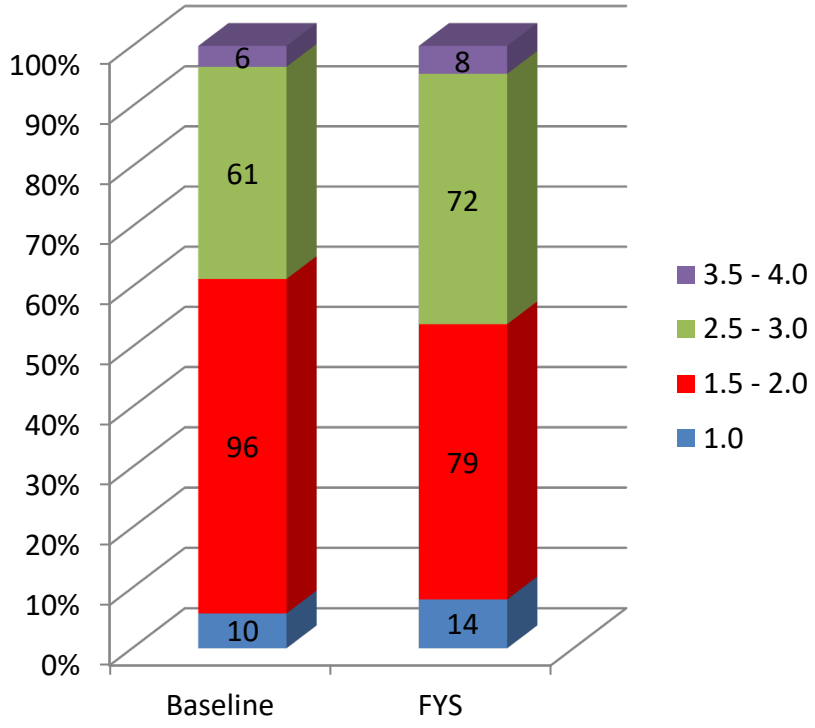
n = 173

Trait/ Performance Level	Evidence Baseline	Evidence FYS	Viewpoints Baseline	Viewpoints FYS	Recommendations Baseline	Recommendations FYS
1.0	18 (10%)	21 (12%)	20 (12%)	24 (14%)	12 (7%)	19 (11%)
1.5 – 2.0	76 (44%)	59 (34%)	122 (71%)	104 (60%)	74 (43%)	50 (29%)
2.5 – 3.0	67 (39%)	84 (49%)	30 (17%)	42 (24%)	73 (42%)	99 (57%)
3.5 – 4.0	12 (7%)	9 (5%)	1 (1%)	3 (2%)	14 (8%)	5 (3%)
Totals	**173**	**173**	**173**	**173**	**173**	**173**

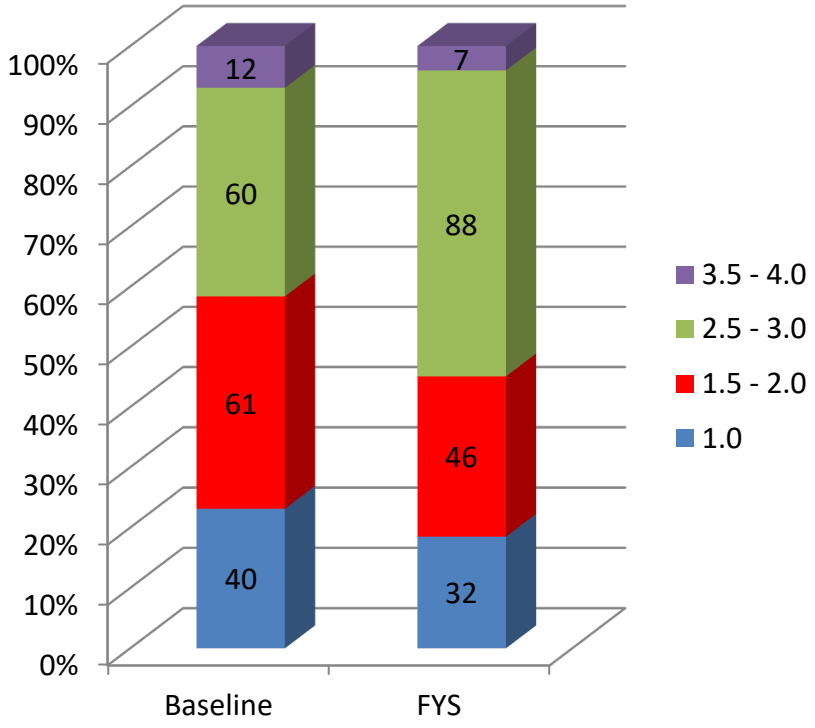
Freshman Baseline/FYS Comparisons

n = 173

Information Needed



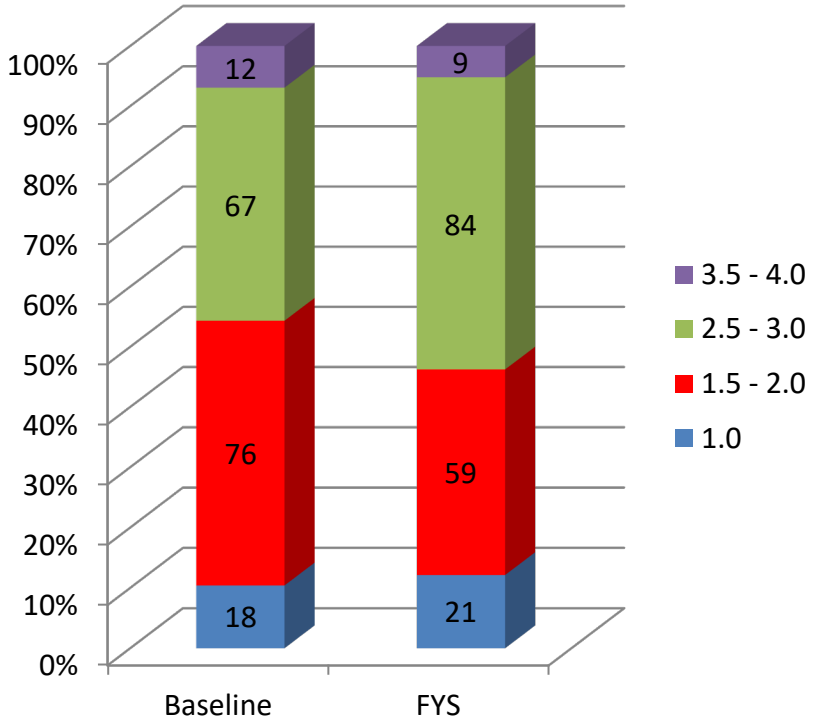
Acknowledgment of Sources



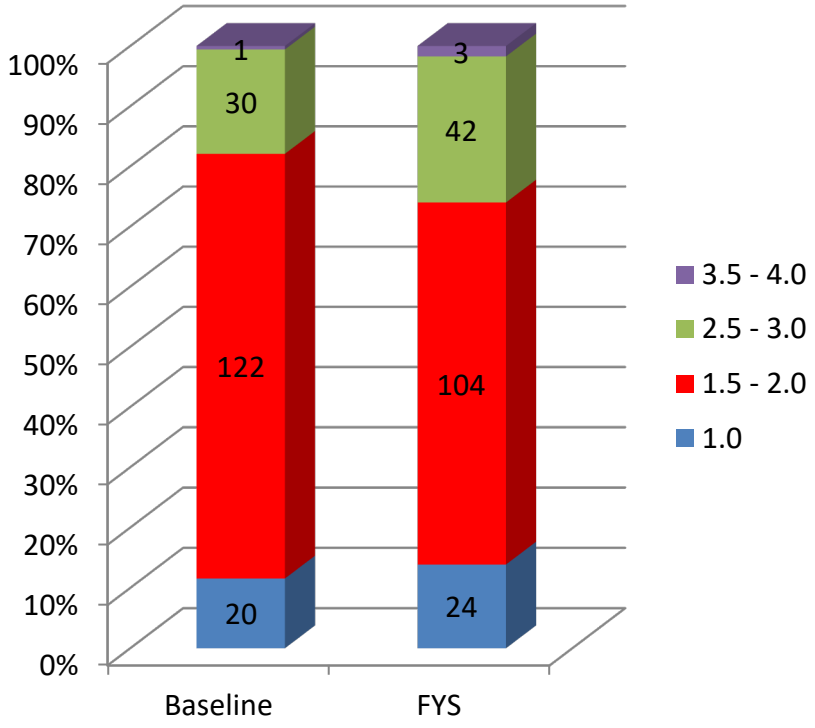
Freshman Baseline/FYS Comparisons

n = 173

Evidence



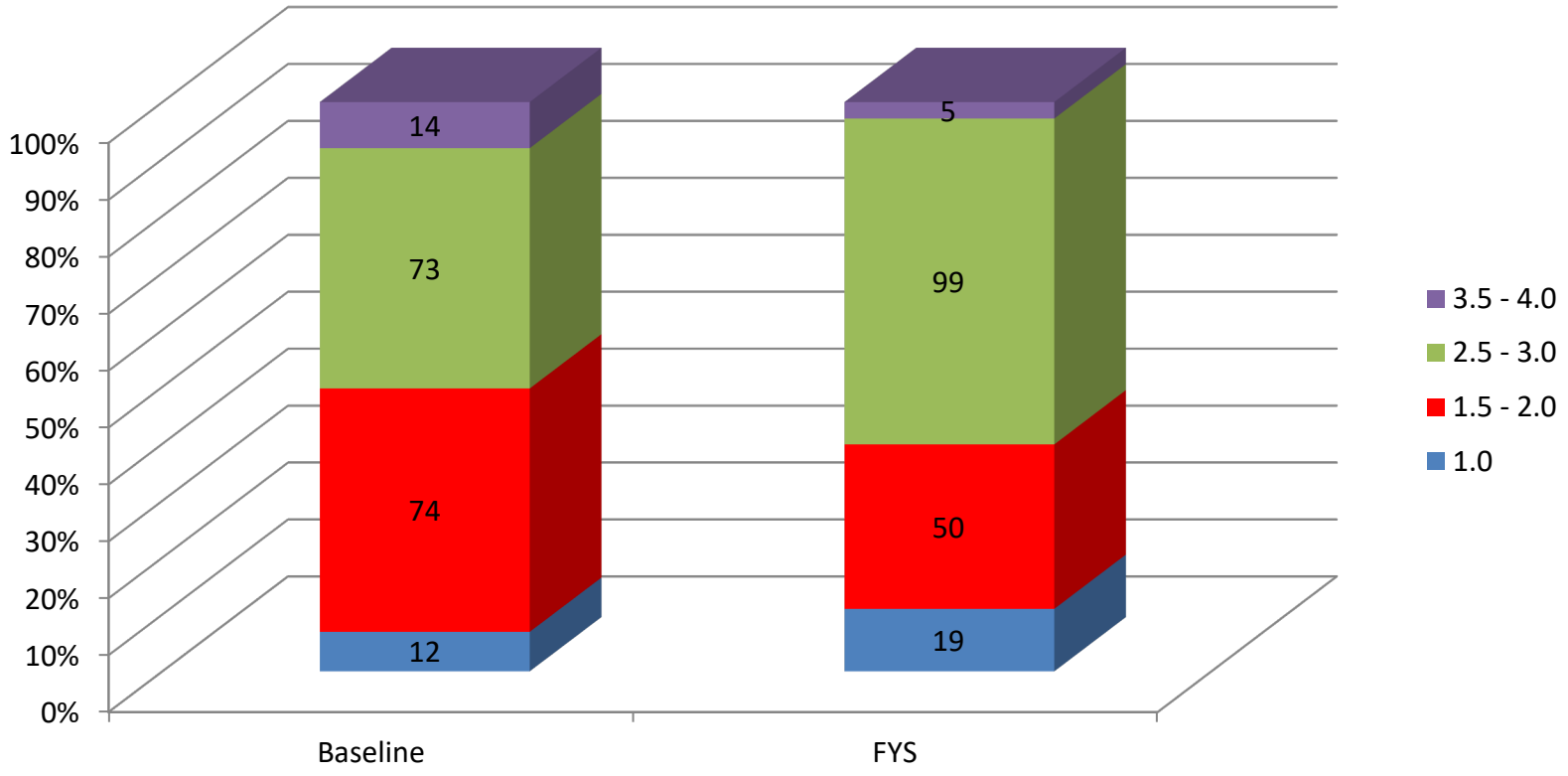
Viewpoints



Freshman Baseline/FYS Comparisons

n = 173

Recommendations



Baseline Inter-Rater Agreement Results

Includes all 175 baseline assessments scored

Trait/ Agreement	Info Needed : Cohen's Kappa (Liberal) = .975	Acknowledgment of Sources: Cohen's Kappa (Liberal) = .880	Evidence: Cohen's Kappa (Liberal) = .958	Viewpoints: Cohen's Kappa (Liberal) = .957	Recommendations: Cohen's Kappa (Liberal) = .936
Agree on score	115 (66%)	100 (57%)	97 (55%)	115 (66%)	93 (53%)
Difference = 1 point	57 (33%)	72 (41%)	72 (41%)	55 (31%)	73 (42%)
Difference = 2 points	3 (2%)	3 (2%)	6 (3%)	5 (3%)	7 (4%)
Difference = 3 points	0	0	0	0	2 (1%)
Total	**175**	**175**	**175**	**175**	**175**

FYS Inter-Rater Agreement Results

Includes all 174 FYS assessments scored

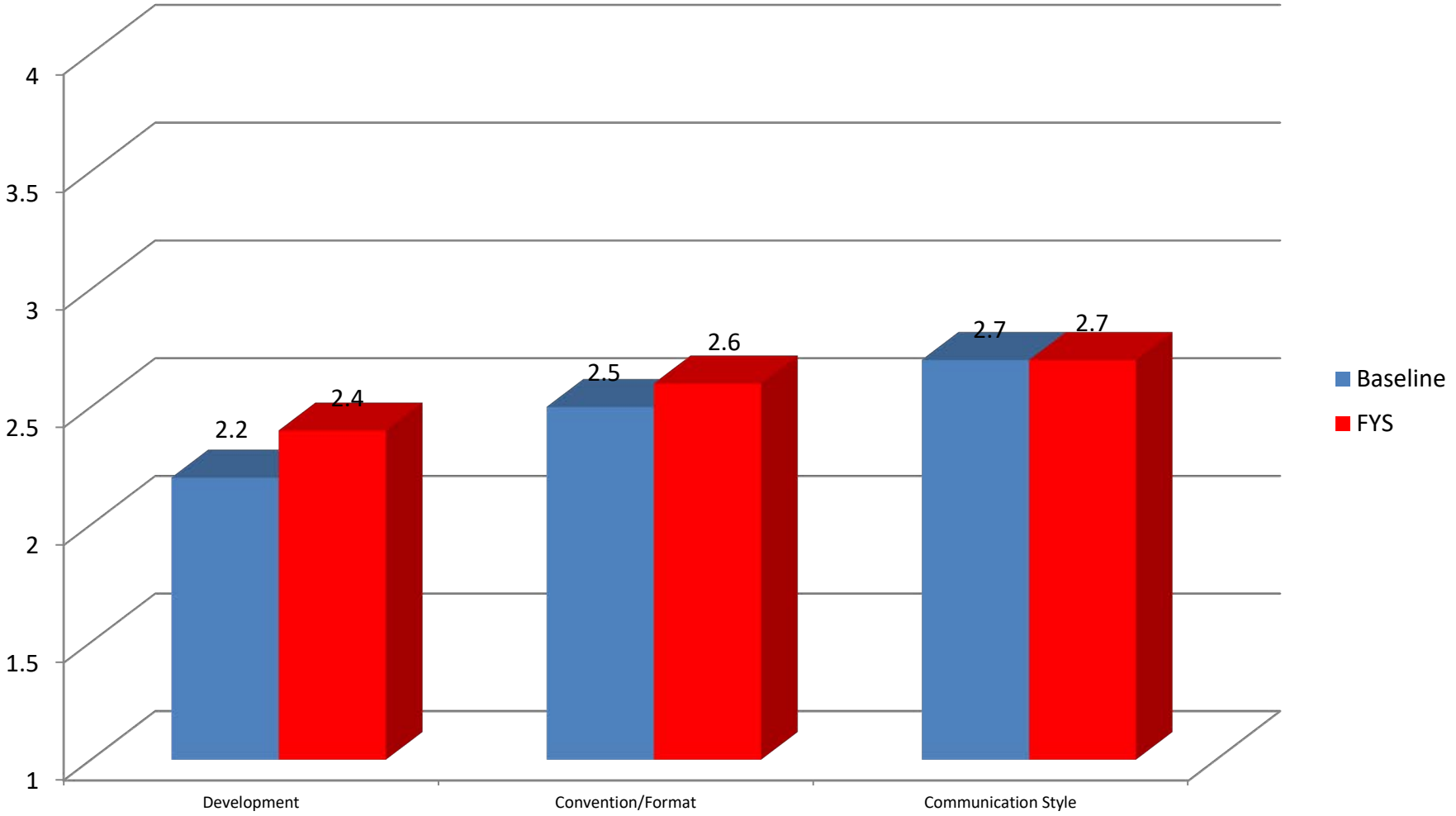
Trait/ Agreement	Info Needed : Cohen's Kappa (Liberal) = .970	Acknowledgment of Sources: Cohen's Kappa (Liberal) = .978	Evidence: Cohen's Kappa (Liberal) = .957	Viewpoints: Cohen's Kappa (Liberal) = .984	Recommendations: Cohen's Kappa (Liberal) = .940
Agree on score	113 (65%)	118 (68%)	95 (55%)	106 (61%)	94 (54%)
Difference = 1 point	57 (33%)	53 (30%)	73 (42%)	66 (38%)	72 (41%)
Difference = 2 points	4 (2%)	3 (2%)	6 (3%)	2 (1%)	8 (5%)
Difference = 3 points	0	0	0	0	0
Total	**174**	**174**	**174**	**174**	**174**

Freshman Baseline/FYS Comparisons

Mean Scores on a scale of 1 – 4, with 4 being the highest possible score

n = 173

Mean differences were statistically significant for *development.*



Freshman Baseline/FYS Comparisons

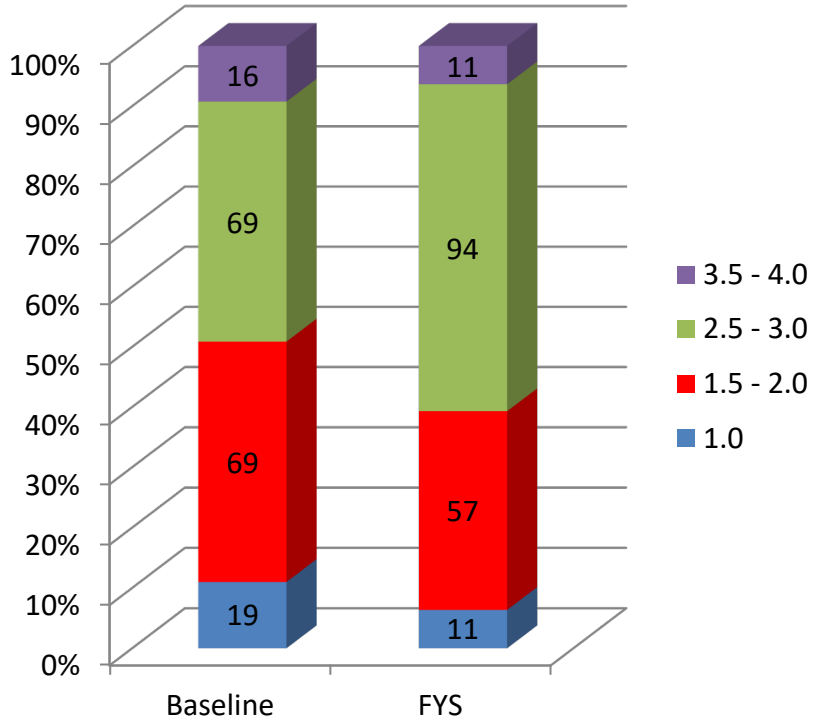
n = 173

Trait/ Performance Level	Development Baseline	Development FYS	Convention/ Format Baseline	Convention/ Format FYS	Communication Style Baseline	Communication Style FYS
1.0	19 (11%)	11 (6%)	11 (6%)	15 (9%)	4 (2%)	3 (2%)
1.5 – 2.0	69 (40%)	57 (33%)	52 (30%)	35 (20%)	30 (17%)	27 (16%)
2.5 – 3.0	69 (40%)	94 (54%)	79 (46%)	96 (55%)	123 (71%)	125 (72%)
3.5 – 4.0	16 (9%)	11 (6%)	31 (18%)	27 (16%)	16 (9%)	18 (10%)
Totals	173	173	173	**173**	**173**	**173**

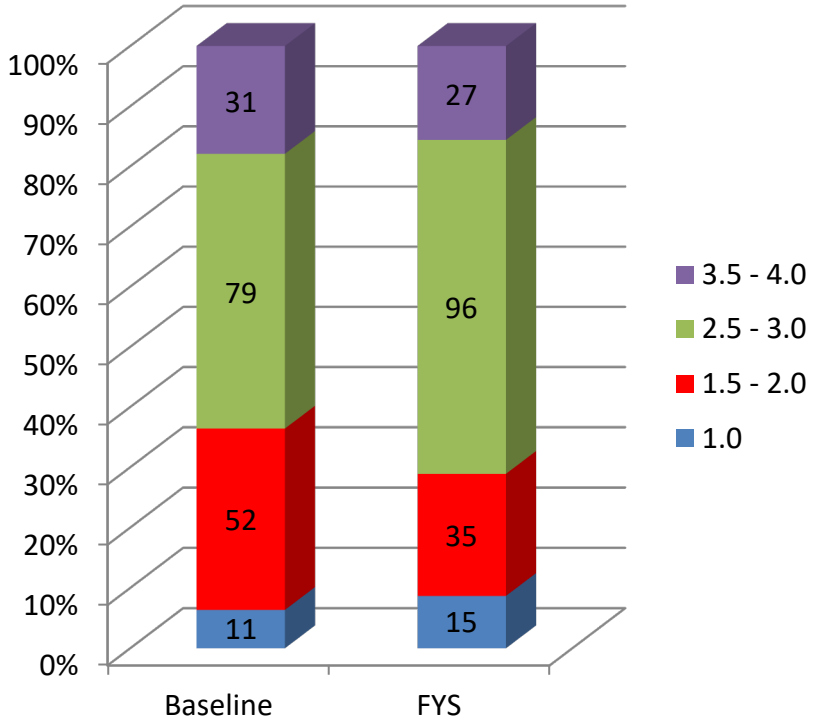
Freshman Baseline/FYS Comparisons

n = 173

Development



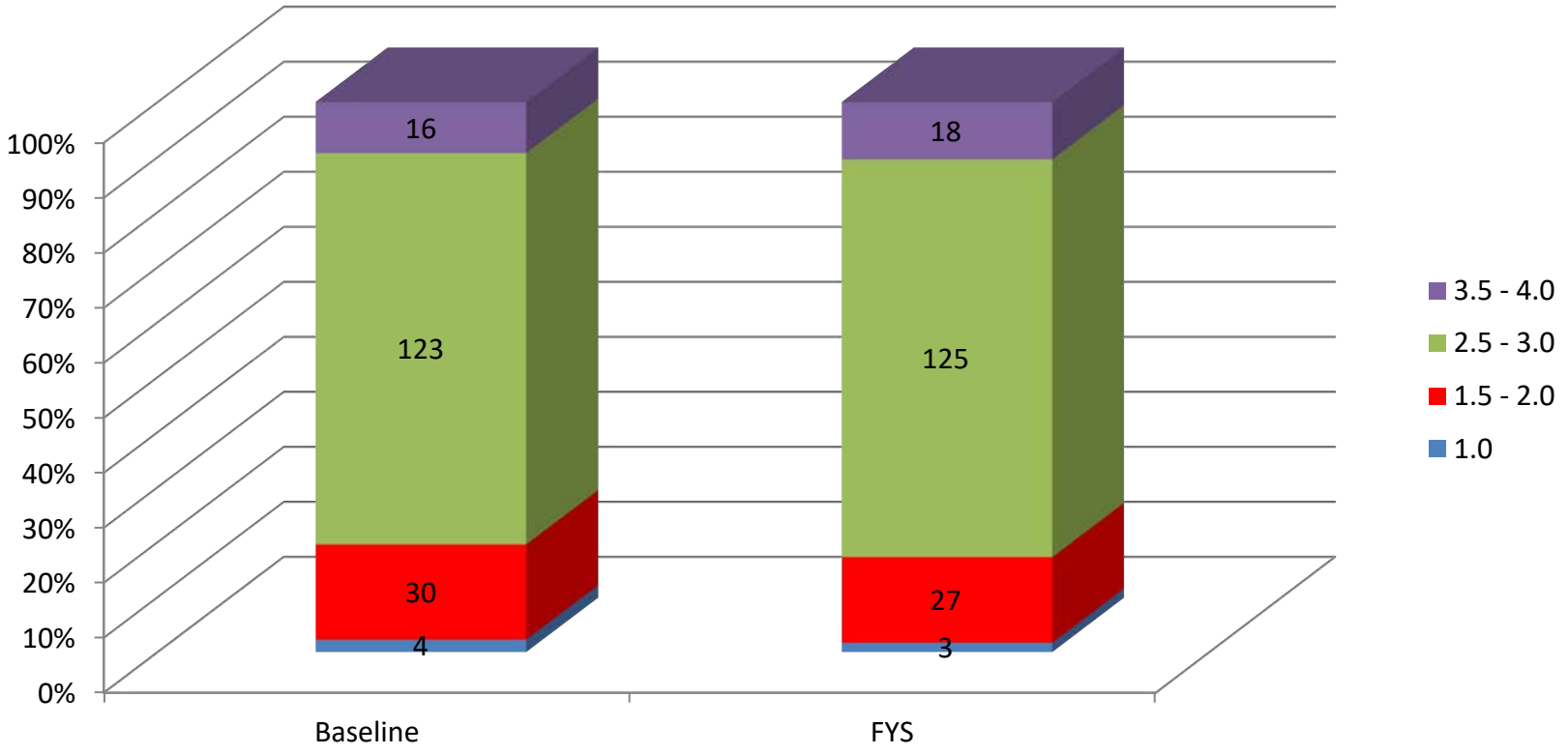
Convention/Format



Freshman Baseline/FYS Comparisons

n = 173

Communication Style



Baseline Inter-Rater Agreement Results

Includes all 175 baseline assessments scored

Trait/ Agreement	Development: Cohen's Kappa (Liberal) = .882	Convention/Format: Cohen's Kappa (Liberal) = .957	Communication Style: Cohen's Kappa (Liberal) = .928
Agree on score	75 (43%)	98 (56%)	102 (58%)
Difference = 1 point	83 (47%)	71 (41%)	65 (37%)
Difference = 2 points	17 (10%)	4 (2%)	7 (4%)
Difference = 3 points	0	2 (1%)	1 (1%)
Total	**175**	**175**	**175**

FYS Inter-Rater Agreement Results

Includes all 174 baseline assessments scored

Trait/ Agreement	Development: Cohen's Kappa (Liberal) = .985	Convention/Format: Cohen's Kappa (Liberal) = .898	Communication Style: Cohen's Kappa (Liberal) = .975
Agree on score	88 (51%)	101 (58%)	105 (60%)
Difference = 1 point	84 (48%)	59 (34%)	66 (38%)
Difference = 2 points	2 (1%)	14 (8%)	3 (2%)
Difference = 3 points	0	0	0
Total	**174**	**174**	**174**

Comparison of FYS Results for Each Trait by Scenario

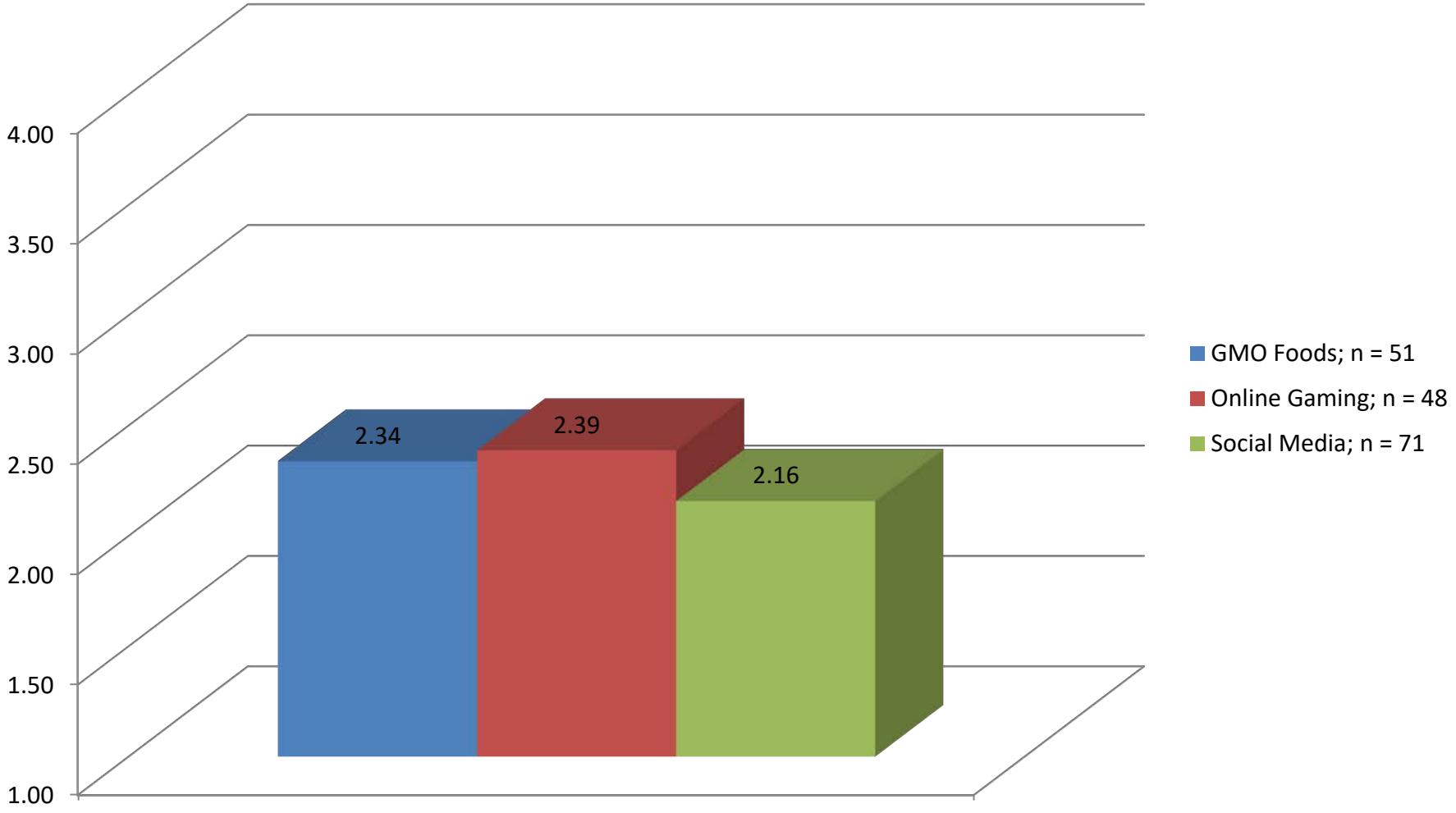
Academic Year 2022 - 2023

FYS Comparisons by Scenario for IL: Information Needed

Mean Scores on a scale of 1 – 4, with 4 being the highest possible score

(The scenarios for two artifacts were unknown and only one aligned to the flu vaccine. These were excluded from this analysis, rendering a final *n* of 170).

A One-Way ANOVA revealed no statistically significant differences in means across the scenarios.

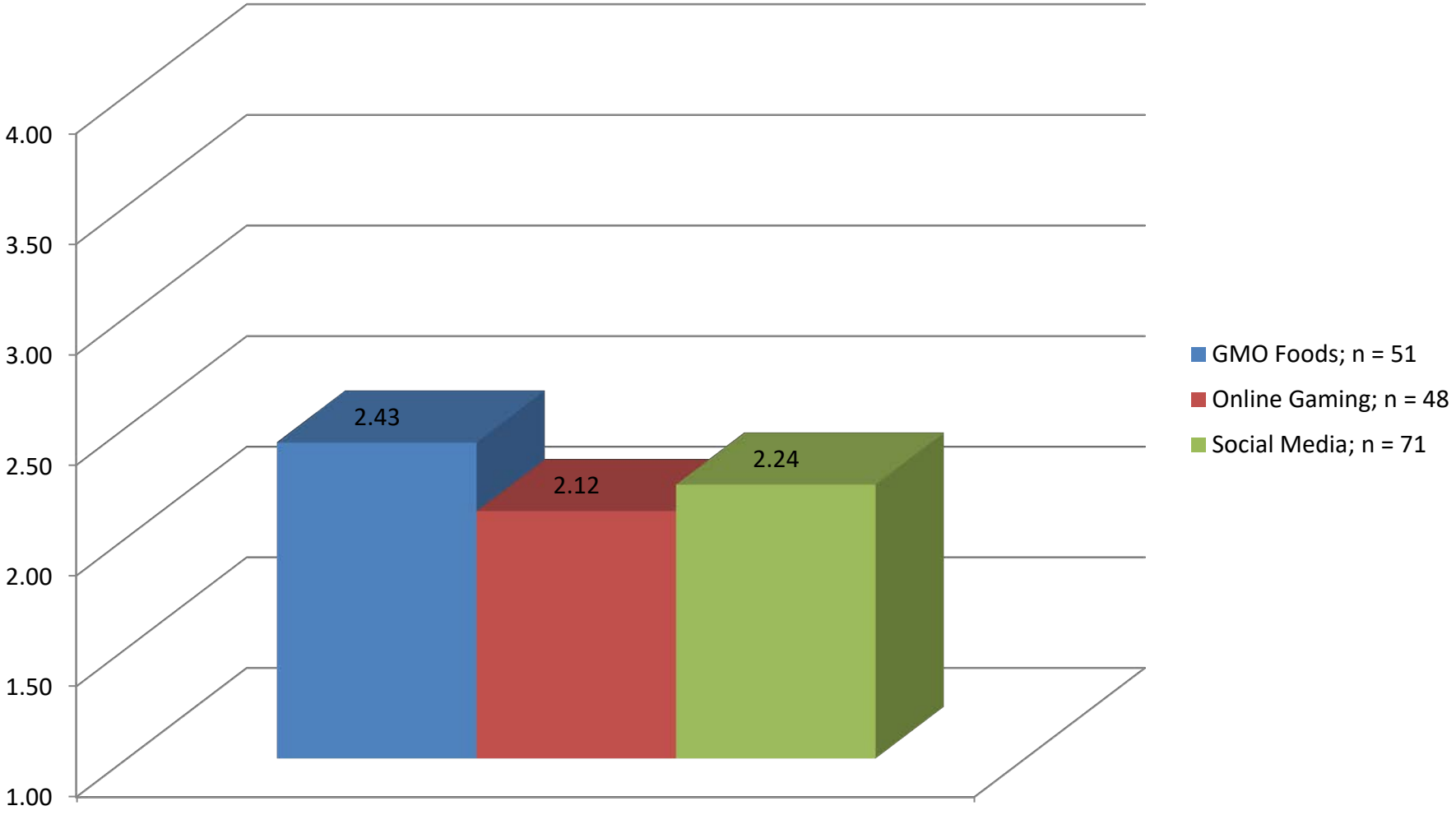


FYS Comparisons by Scenario for IL: Source Acknowledgment

Mean Scores on a scale of 1 – 4, with 4 being the highest possible score

(The scenarios for two artifacts were unknown and only one aligned to the flu vaccine. These were excluded from this analysis, rendering a final *n* of 170).

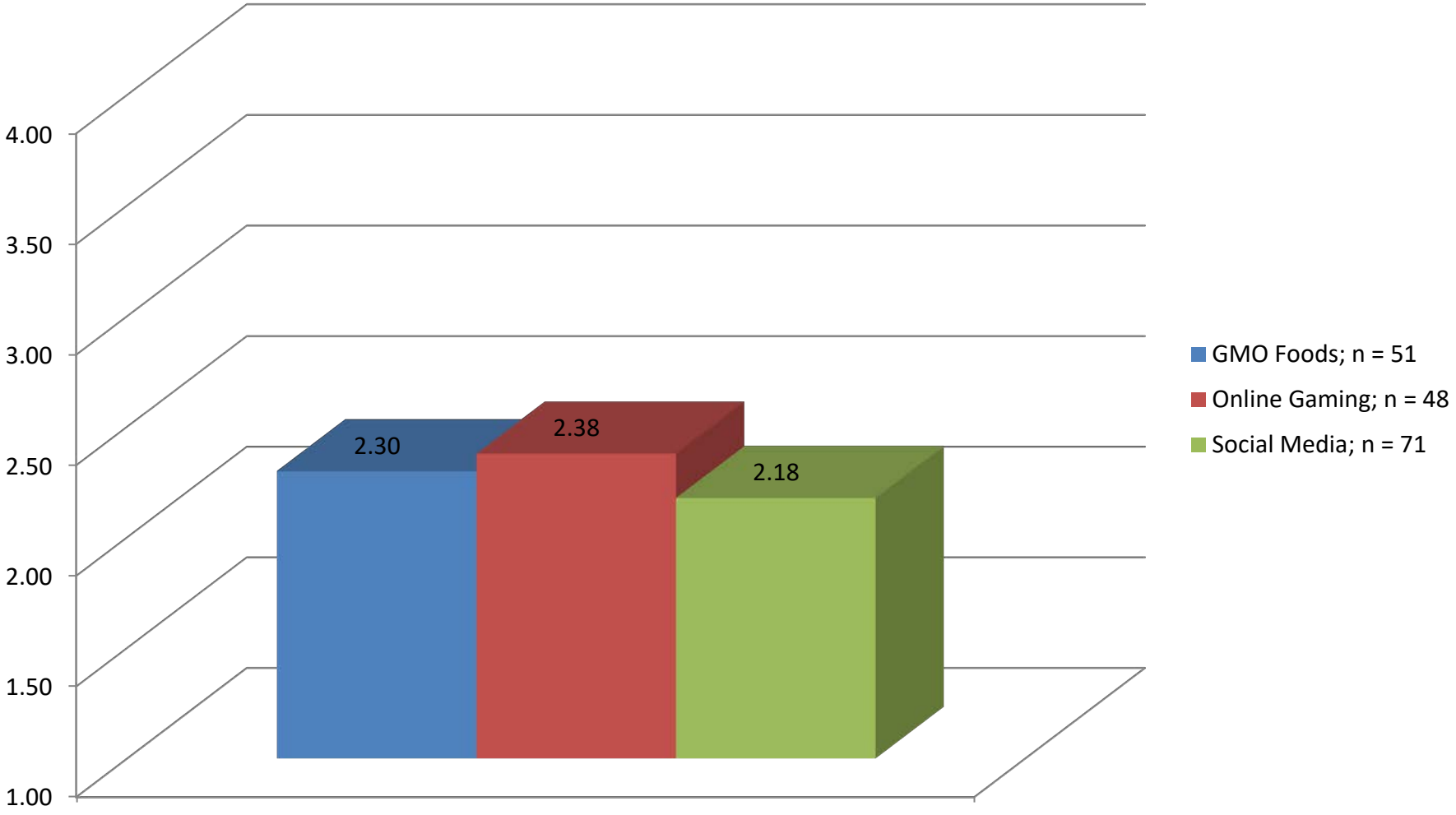
A One-Way ANOVA revealed no statistically significant differences in means across the scenarios.



FYS Comparisons by Scenario for CT: Evidence

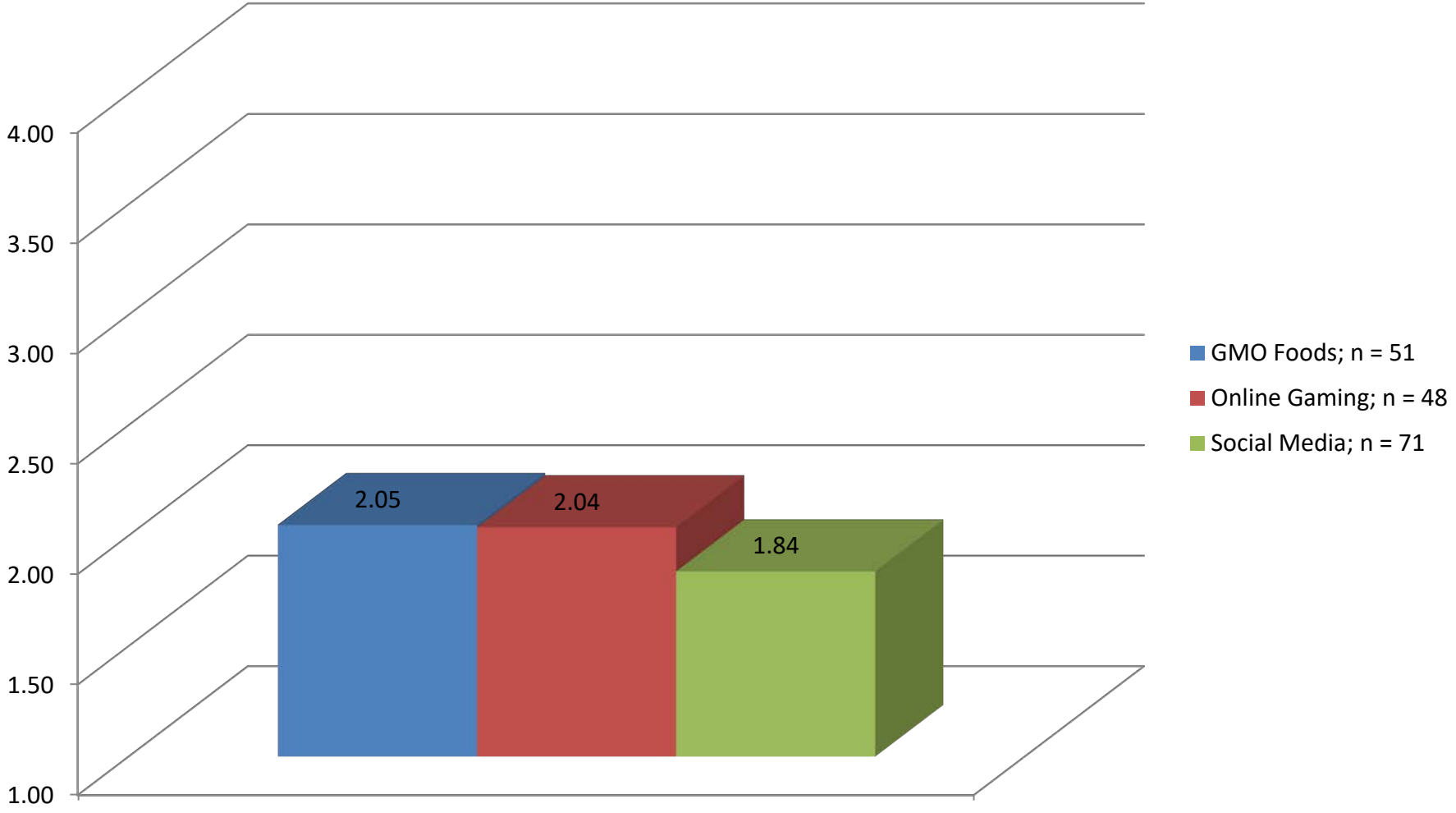
Mean Scores on a scale of 1 – 4, with 4 being the highest possible score
(The scenarios for two artifacts were unknown and only one aligned to the flu vaccine. These were excluded from this analysis, rendering a final *n* of 170).

A One-Way ANOVA revealed no statistically significant differences in means across the scenarios.



FYS Comparisons by Scenario for CT: Viewpoints

Mean Scores on a scale of 1 – 4, with 4 being the highest possible score
(The scenarios for two artifacts were unknown and only one aligned to the flu vaccine. These were excluded from this analysis, rendering a final *n* of 170).
A One-Way ANOVA revealed no statistically significant differences in means across the scenarios.

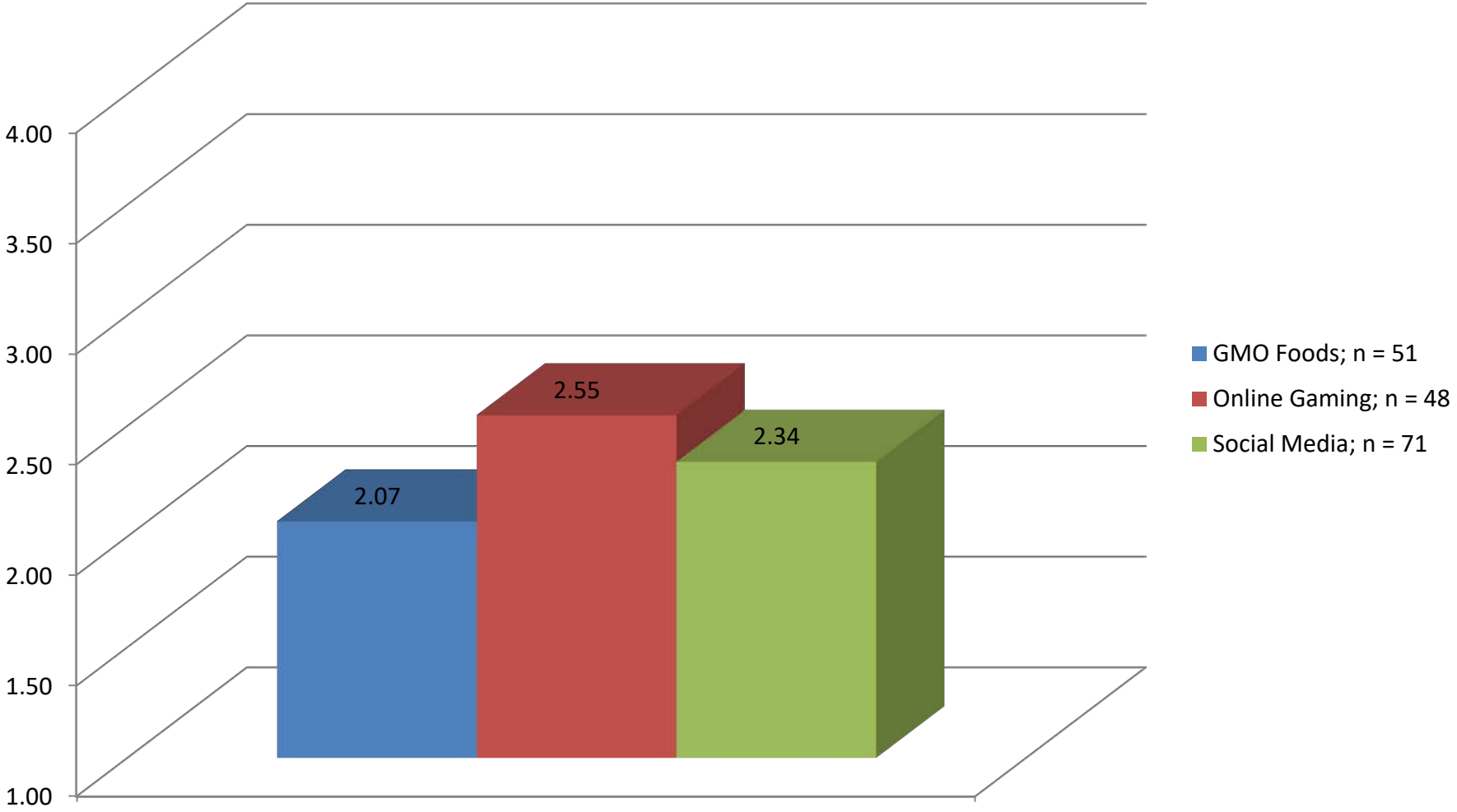


FYS Comparisons by Scenario for CT: Recommendation/Position

Mean Scores on a scale of 1 – 4, with 4 being the highest possible score

(The scenarios for two artifacts were unknown and only one aligned to the flu vaccine. These were excluded from this analysis, rendering a final *n* of 170).

A One-Way ANOVA showed statistical significance across the scenarios. Bonferroni post-hoc analysis revealed that the mean for GMO Foods was significantly lower than the mean for Online Gaming.

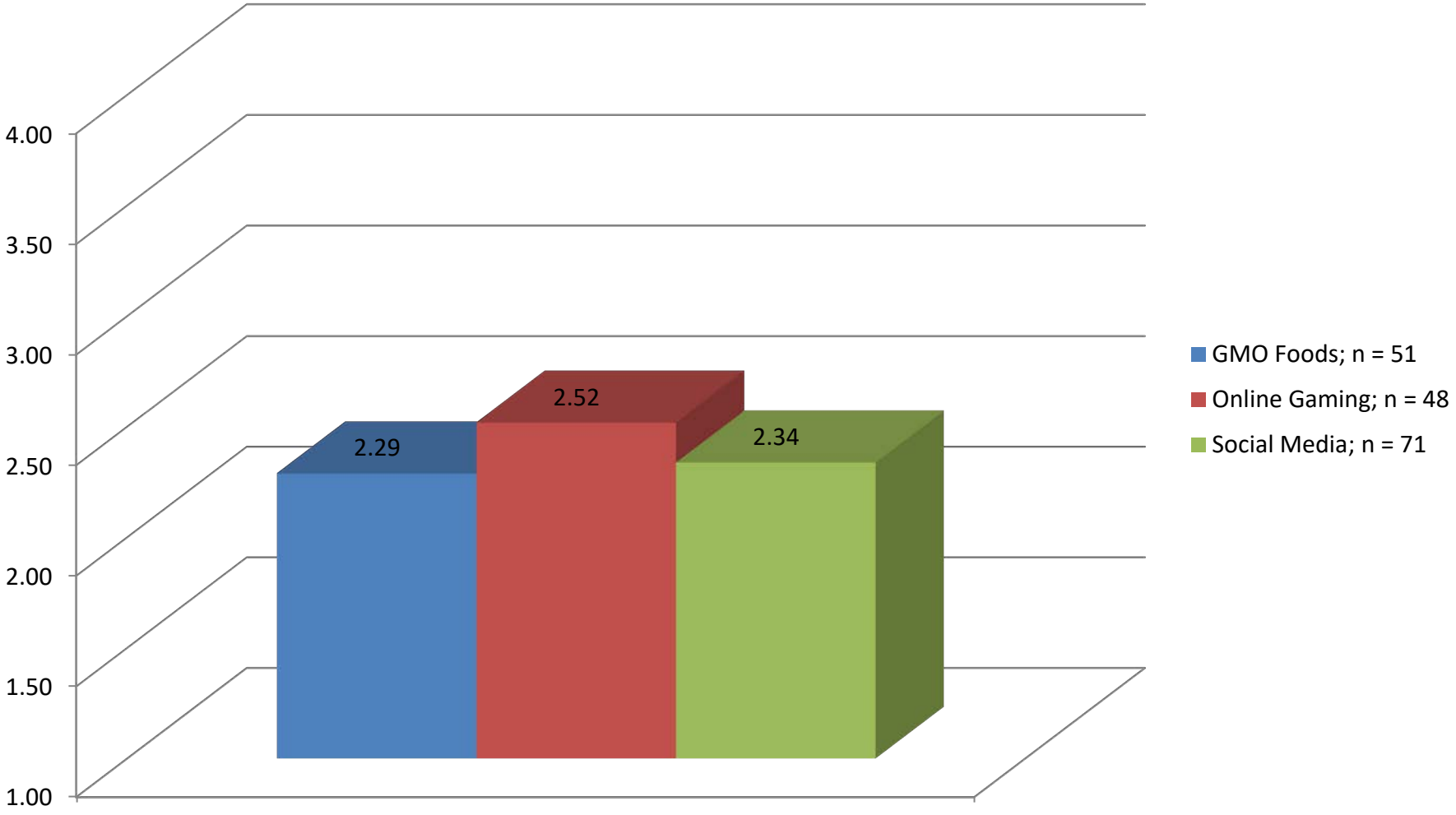


FYS Comparisons by Scenario for CF: Development

Mean Scores on a scale of 1 – 4, with 4 being the highest possible score

(The scenarios for two artifacts were unknown and only one aligned to the flu vaccine. These were excluded from this analysis, rendering a final *n* of 170).

A One-Way ANOVA revealed no statistically significant differences in means across the scenarios.

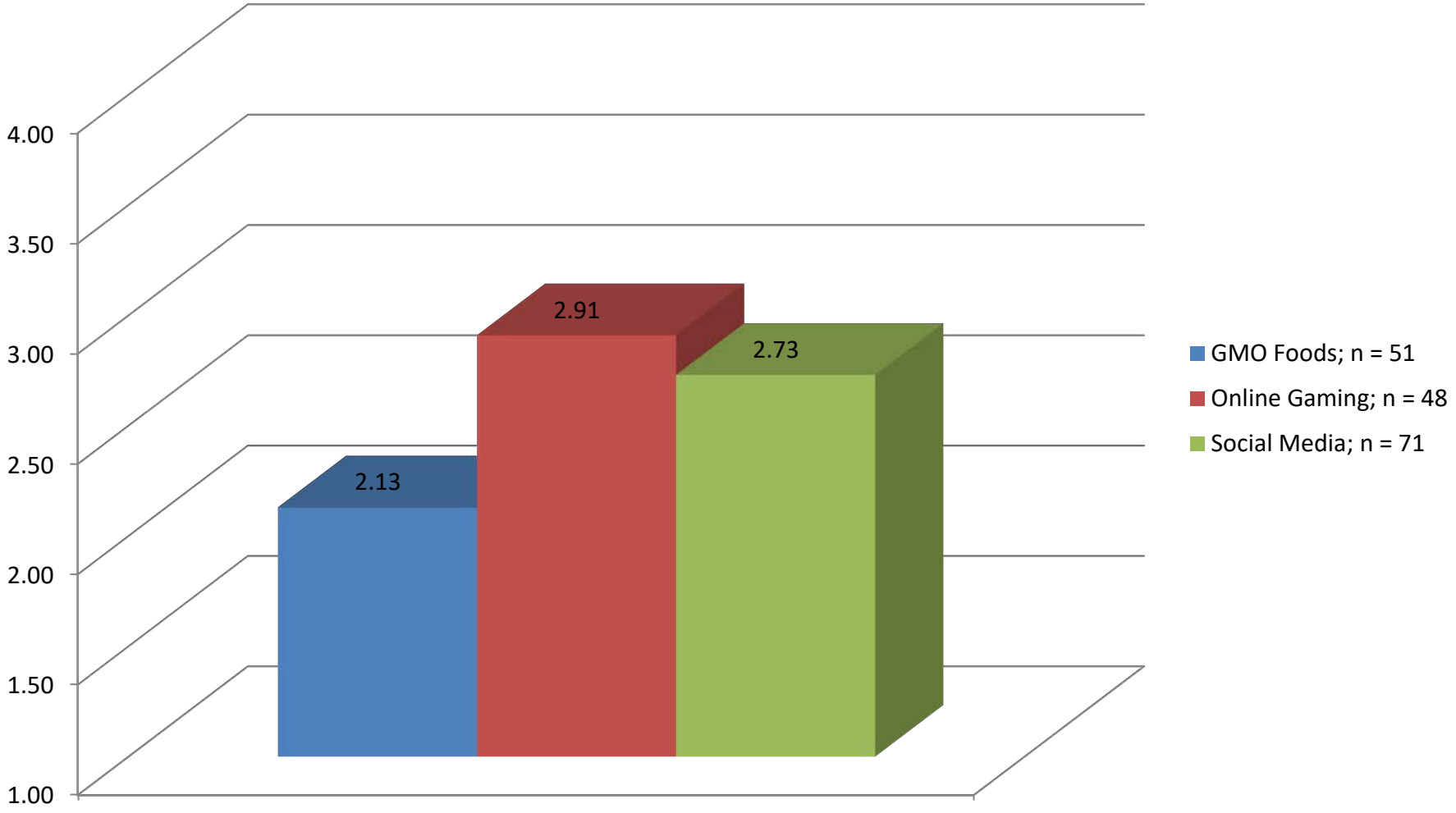


FYS Comparisons by Scenario for CF: Convention/Format

Mean Scores on a scale of 1 – 4, with 4 being the highest possible score

(The scenarios for two artifacts were unknown and only one aligned to the flu vaccine. These were excluded from this analysis, rendering a final *n* of 170).

A One-Way ANOVA revealed significant mean differences among scenarios; Bonferroni post-hoc analysis revealed that the mean for GMO Foods was significantly lower than the means for Online Gaming and Social Media.

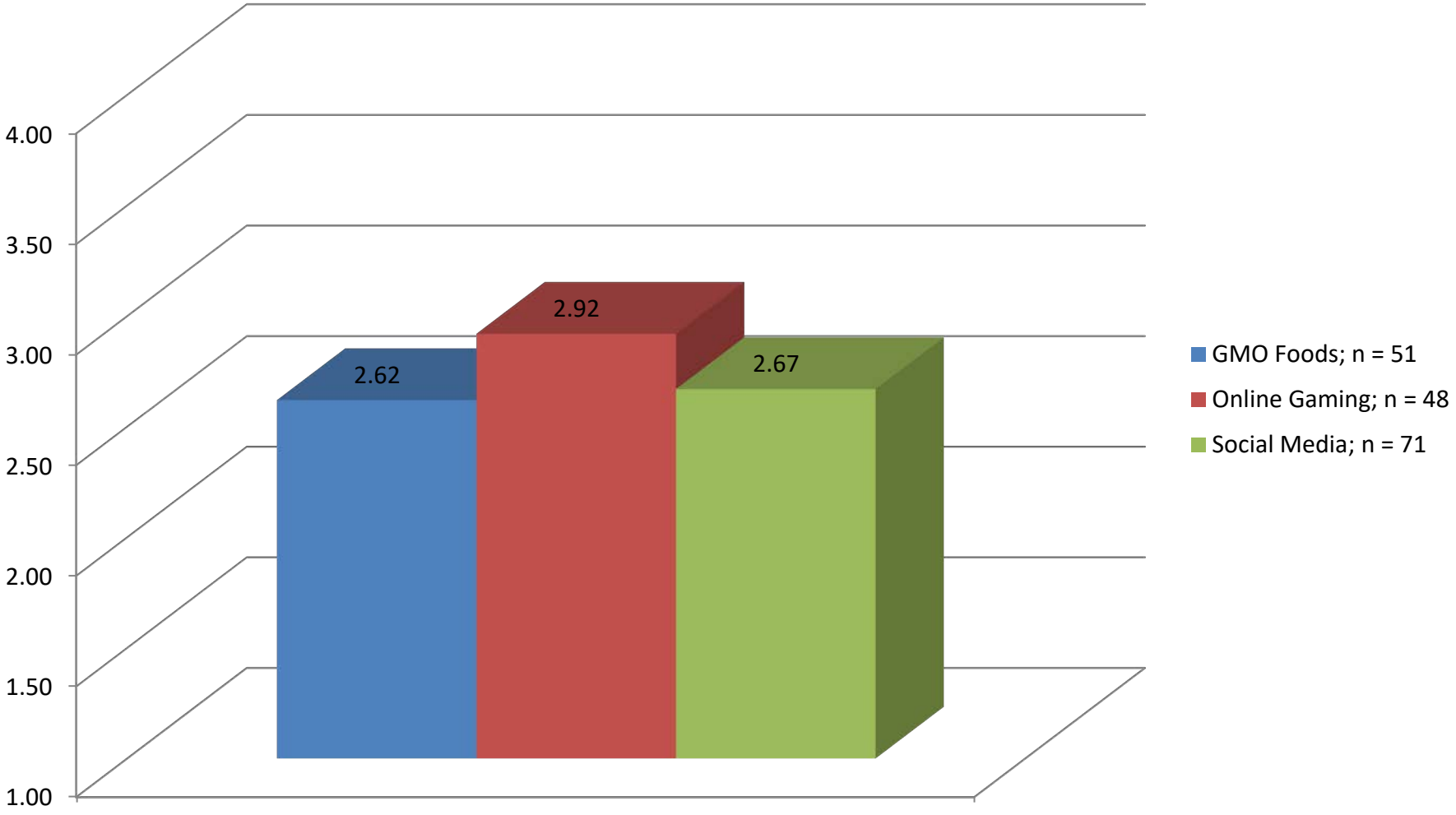


FYS Comparisons by Scenario for CF: Communication Style

Mean Scores on a scale of 1 – 4, with 4 being the highest possible score

(The scenarios for two artifacts were unknown and only one aligned to the flu vaccine. These were excluded from this analysis, rendering a final *n* of 170).

A One-Way ANOVA revealed statistical significance across scenarios; Bonferroni post-hoc analysis revealed small significant mean differences between GMO Foods and Online Gaming and between Online Gaming and Social Media.



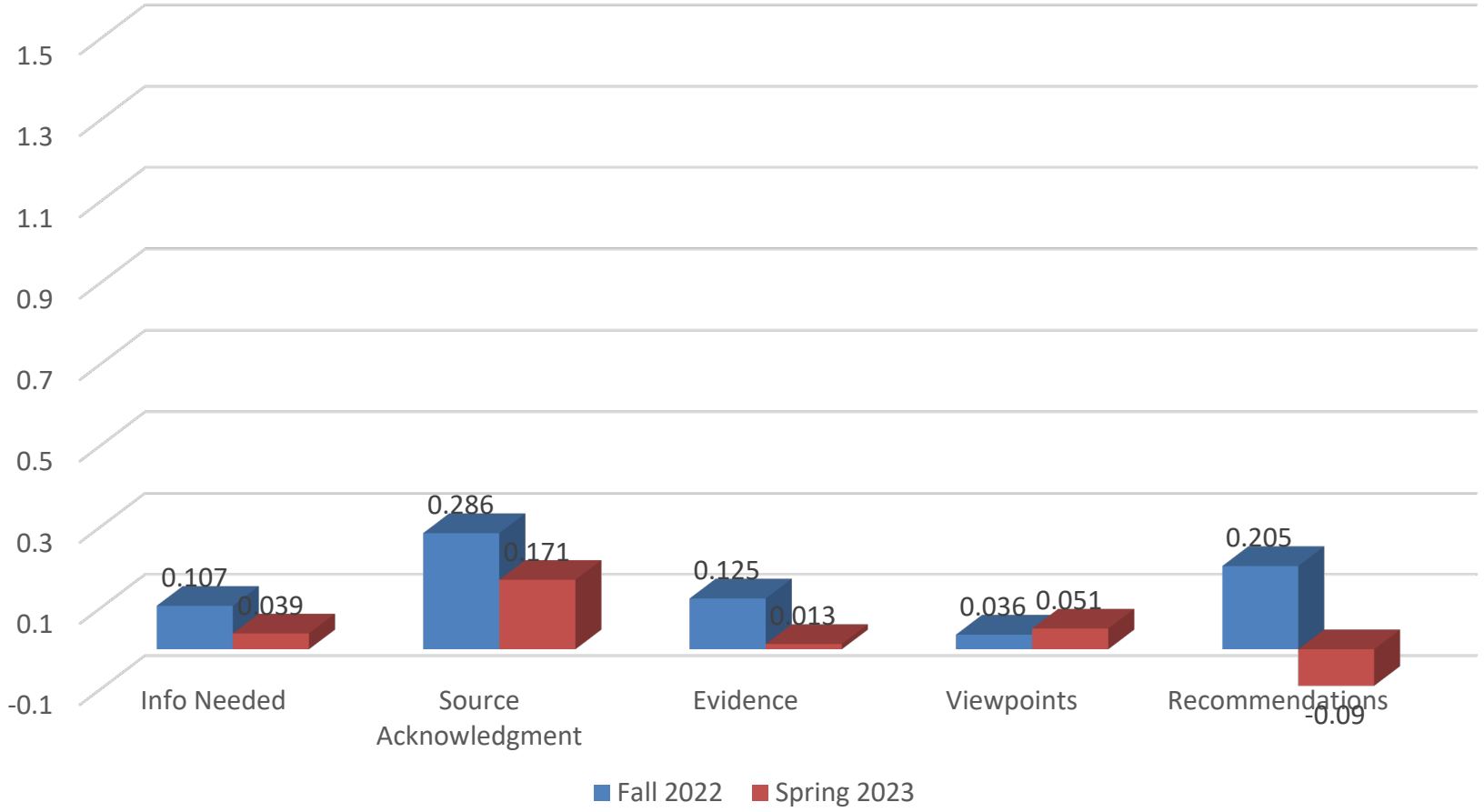
Comparison of Baseline to FYS Mean Gain Score for Each Trait by Semester of FYS

Academic Year 2022 - 2023

Baseline to FYS Mean Gain Scores for Each Trait

n = 56 in fall and 117 in spring

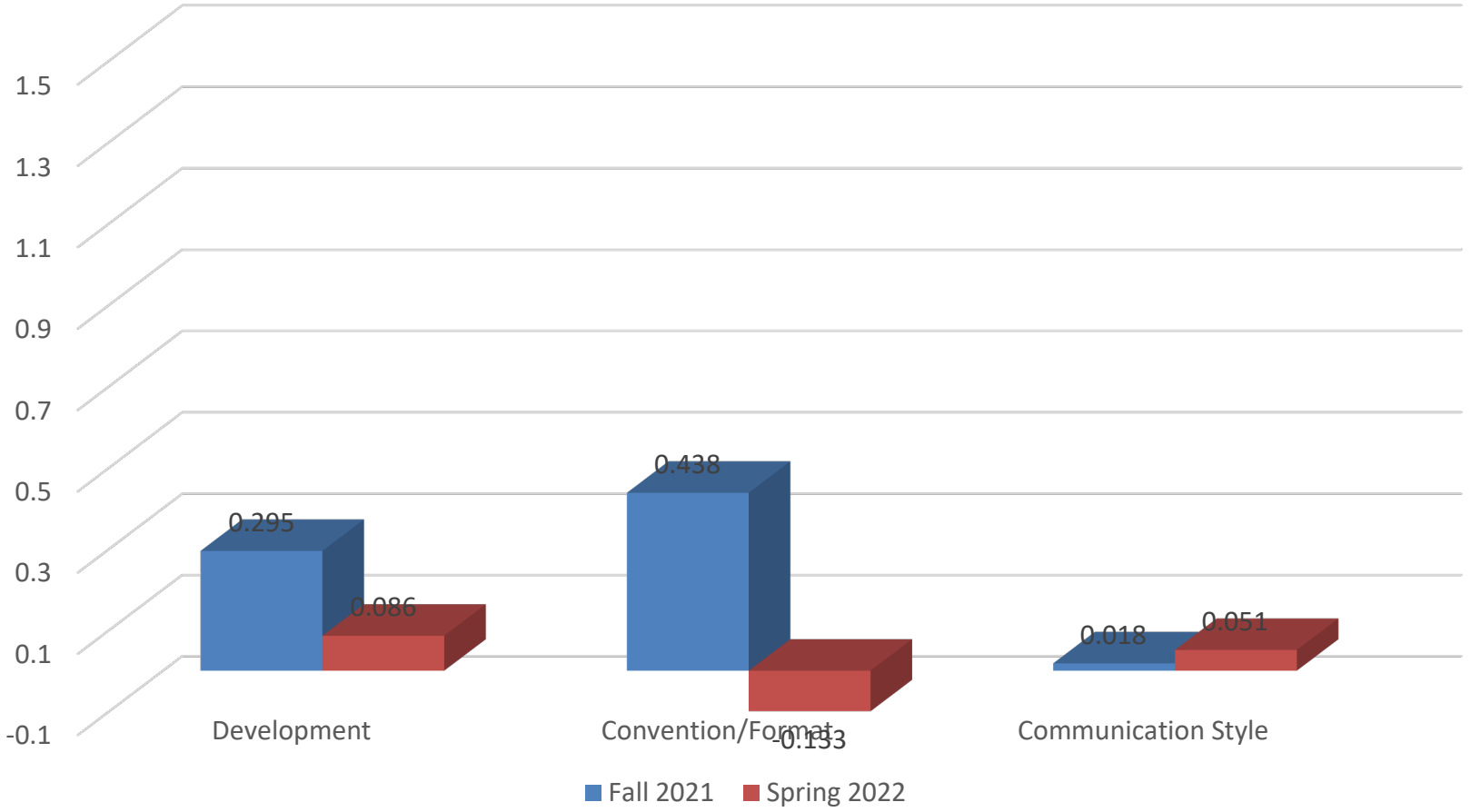
(Mean differences between fall and spring were not statistically significant)



Baseline to FYS Mean Gain Scores for Each Trait

n = 56 in fall and 117 in spring

(Mean difference between fall and spring for *Convention/Format* were significant, *t* (110.183) = -3.813; *p* < .001. *Mean differences for the other traits were not significant*).



Reference

Stellmack, M.A., Kohneim-Kalkstein, Y. L, Manor, J. E., Massey, A. R., & Schmitz, J. A. P. (2009). An assessment of reliability and validity of a rubric for grading APA-style introductions. *Teaching of Psychology*, *36*, 102-107.