

# Assessment Report

October 2015

## 1 Introduction

During the summer of 2015, three groups of graders met to grade artifacts using rubrics created to measure the Written Communication, Critical and Creative Thinking and Research Fluency general education outcomes. Data was collected using a stratified sampling method in which two artifacts were selected at random from each section of First Year Seminar (FYS) and First Year Writing (FYW). Artifacts were taken from courses offered only during the spring semester in order to ensure that each student had had a comparable amount of writing instruction. Each group attempted to grade 30 FYS and 30 FYW artifacts.

The purpose of collecting the data was two-fold. First, we wanted to test the validity of the rubric by measuring inter-rater reliability. To this end, each artifact was graded independently by two raters on a scale of 0 to 4. In addition, we wanted to determine how well our students are meeting the general education outcomes by looking at the rubric scores themselves. As each artifact was graded twice, the individual rubric scores are not independent so it would be misleading to treat them as such; thus the median score for each rubric item is reported.

In order to be able to compare 2014 to 2015 results, a sample of artifacts from 2014 were regraded and a hypothesis test was performed to see if there was any significant difference in grades assigned. In all cases it was determined that there was no significant difference in grades assigned and therefore any change in median scores should not be attributed to variation in the grader groups.

The report is organized by general education outcome. For each outcome a summary of student performance and agreement data for all rubric items is provided followed by more detailed summaries of individual rubric items. The individual rubric item summaries include stacked bar graphs comparing 2014 to 2015 data, single bar graphs displaying combined data for two years, and agreement data categorized by type of disagreement.

As in 2014, agreement data is reported using both percent agreement and Gwet's AC2 statistic. One advantage to using Gwet's AC2 statistic is that we are able to capture both agreement, the percentage of artifacts on which the two raters assign the same score, and near agreement, the percentage of artifacts on which the two raters differ by one point or less. Generally when a four point scale is used, only percent agreement is reported, since near agreement accounts for ten of the sixteen ways two raters could assign scores. The minimum acceptable standard for percent agreement on a four point scale is 70%. Another advantage to using Gwet's AC2 statistic is that it is a chance corrected agreement statistic. If two raters were to randomly assign scores using a four point scale, the probability that they would agree by chance is  $4/16=0.25$ . When percent agreement is measured, it is not clear whether the raters actually agreed on the score, or the agreement was due to chance. One interpretation of Gwet's AC2 statistic is given in the table below:

AC2	Interpretation
< .20	Poor
.21-.40	Fair
.41-.60	Moderate
.61-.80	Good
.81-1.00	Very Good

Table 1: Interpretation of Gwet's AC2 Statistic

## 2 Written Communication

### 2.1 Summary of the Written Communication Outcome

With the exception of the Sources and Evidence rubric item, our students tended to perform better in 2014 than in 2015 on the Written Communication outcome. However, as seen in the bar chart below, the Sources and Evidence rubric item had a high percentage of low scores (median scores of 0 or 0.5) in both 2014 and 2015.

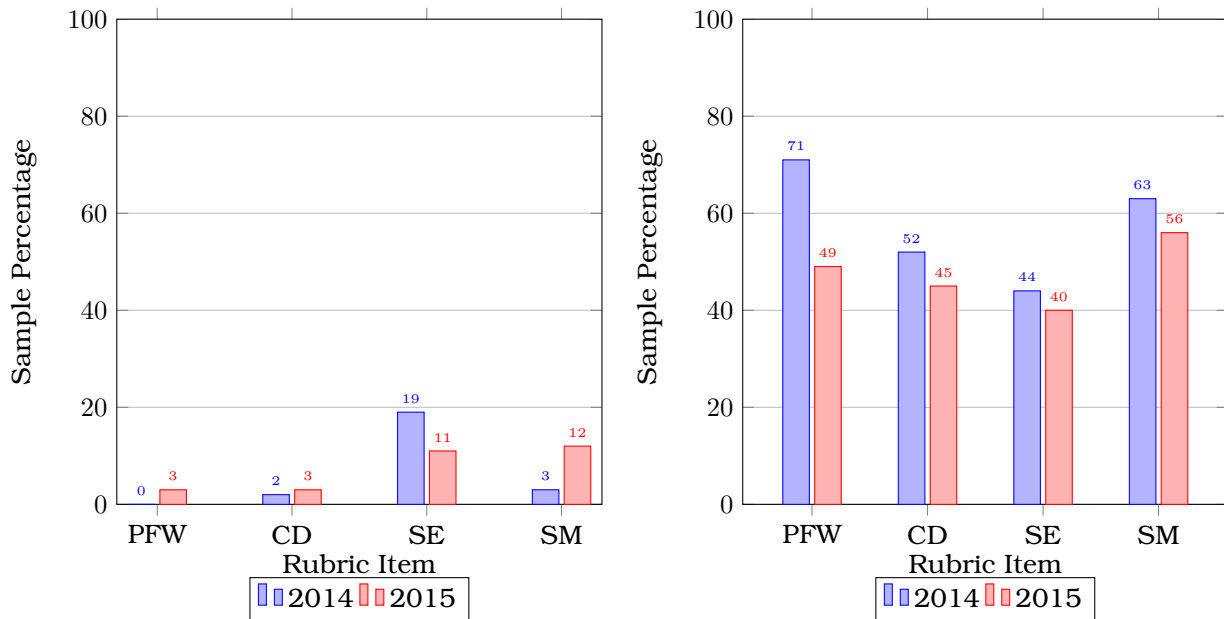


Figure 1: The figure on the left shows the percentage of the sample with  $0 \leq \text{median score} \leq 0.5$ , while the one on the right shows the percentage with  $2 \leq \text{median score} \leq 3$ .

All of the Written Communication rubric items had at least moderate agreement in 2015. The only rubric items showing change in the value of the AC2 statistic were the Sources and Evidence rubric item which decreased slightly but remained at the moderate level and the Control of Syntax and Mechanics rubric item which increased from fair to good agreement.

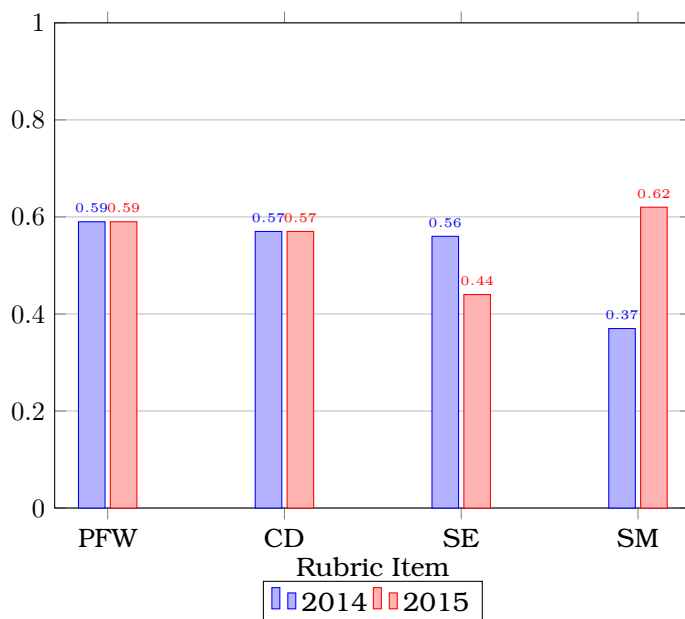


Figure 2: Gwet's AC2 - Written Communication

The data gathered in this sample can also be used to draw conclusions about all Rhode Island College freshmen. The tables below give 95% confidence intervals for the percentage of RIC freshmen obtaining low and high median scores, respectively. Focusing, for example, on the Purpose for Writing rubric item, we could say that we are 95% confident that between 0% and 7% of RIC freshmen would receive a median score of 0 or 0.5, while between 39% and 59% of RIC freshmen would receive a median score between 2 and 3 inclusive. Similar interpretations would apply to the other rubric items.

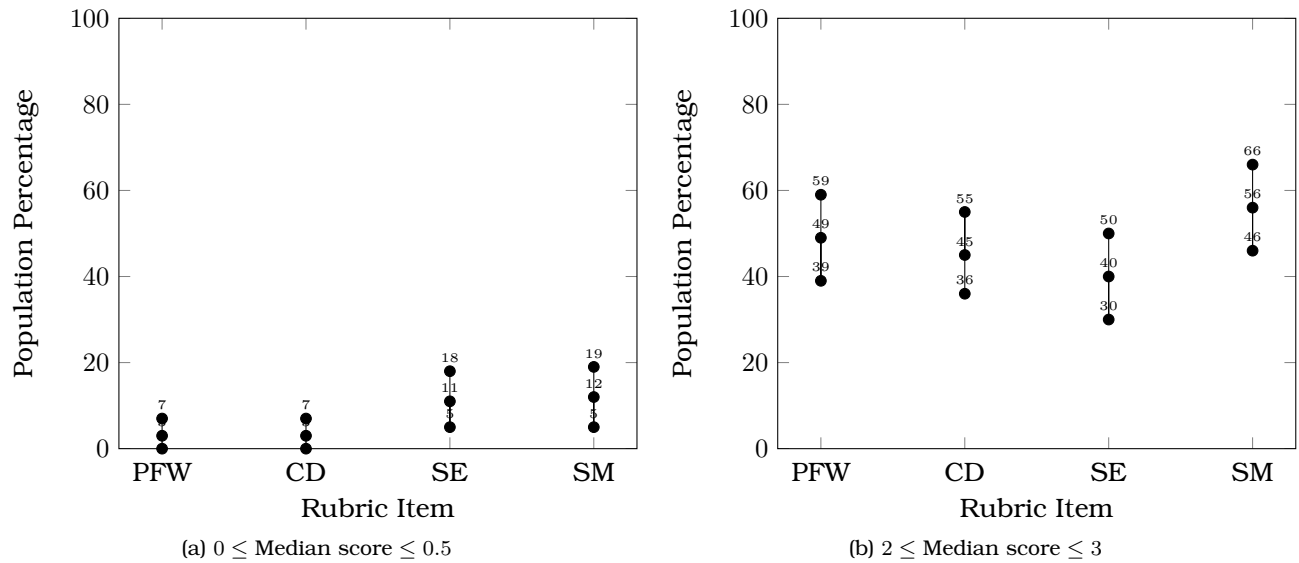


Figure 3: 95% Confidence Intervals for the Population Percentage

## 2.2 Written Communication Summaries by Rubric Item

### 2.2.1 Student Performance

The combined two year data charts that follow for each rubric item indicate that the distribution of scores for the Purpose for Writing, and Control of Syntax and Mechanics rubric items is skewed slightly to the left meaning these rubric items had more high than low median scores.

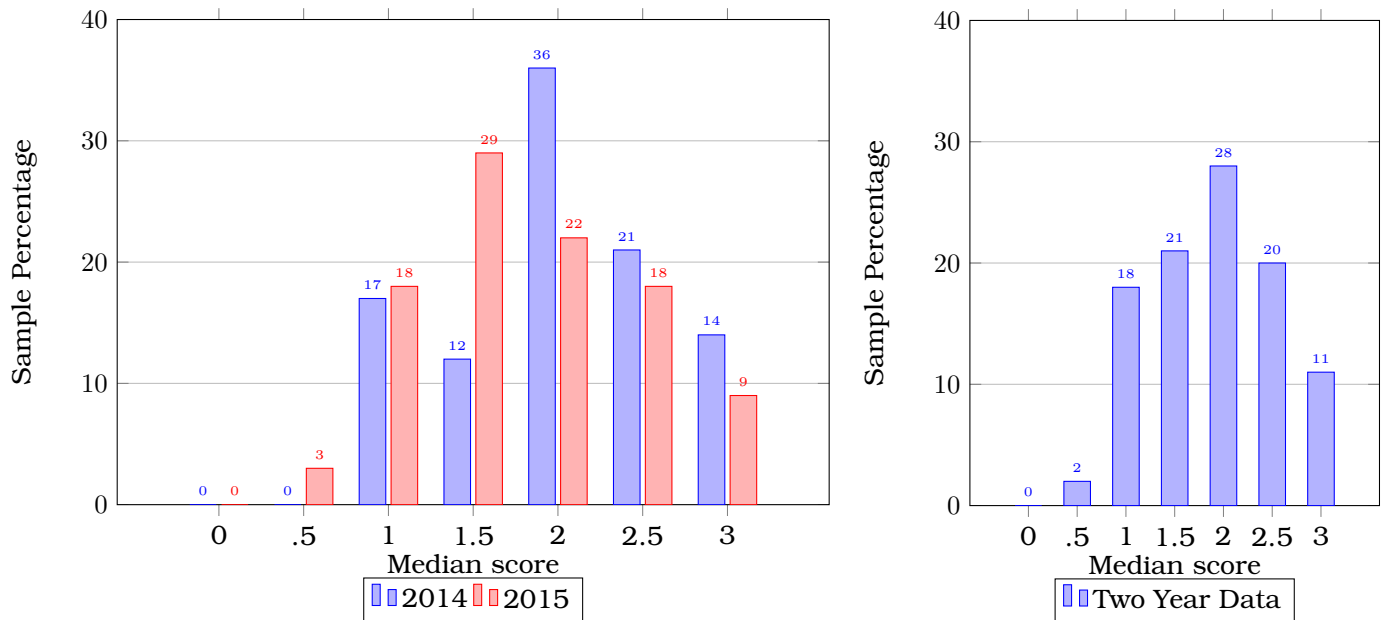


Figure 4: Purpose for Writing

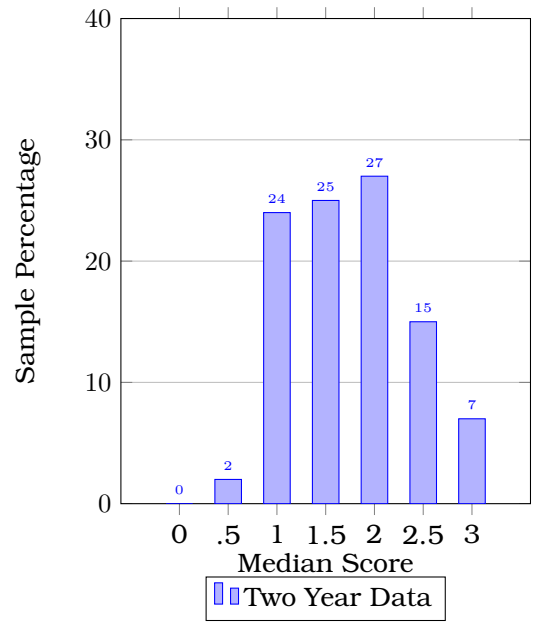
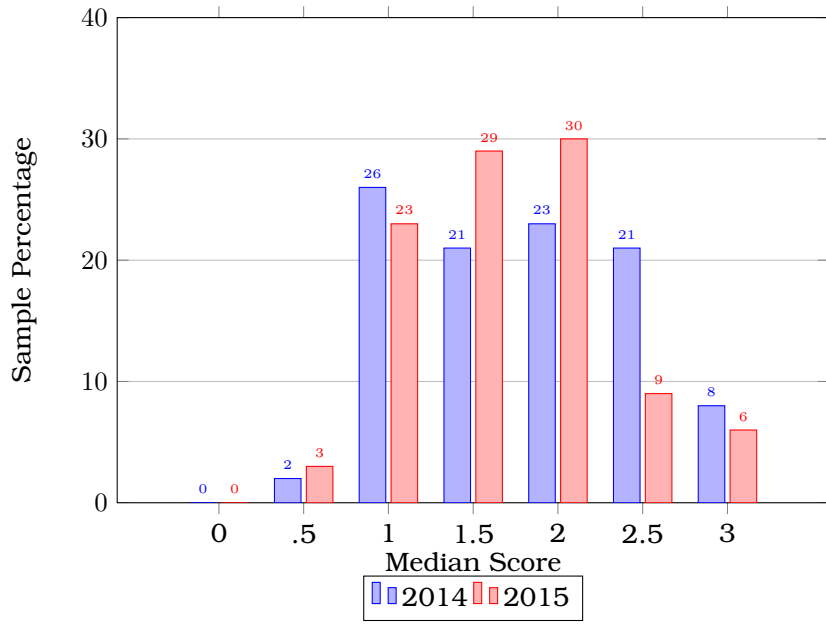


Figure 5: Content Development

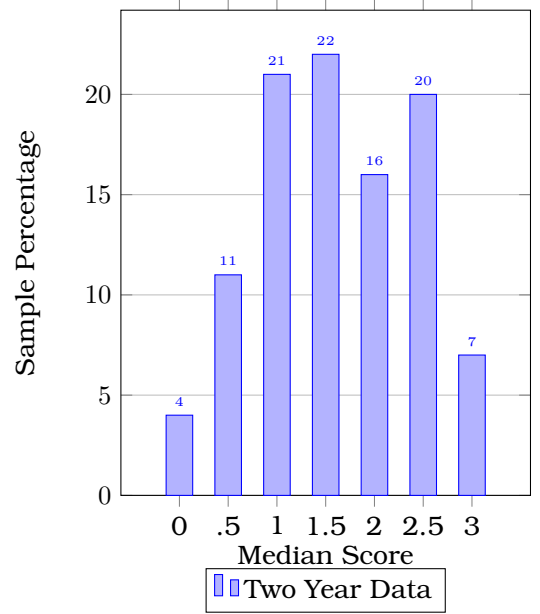
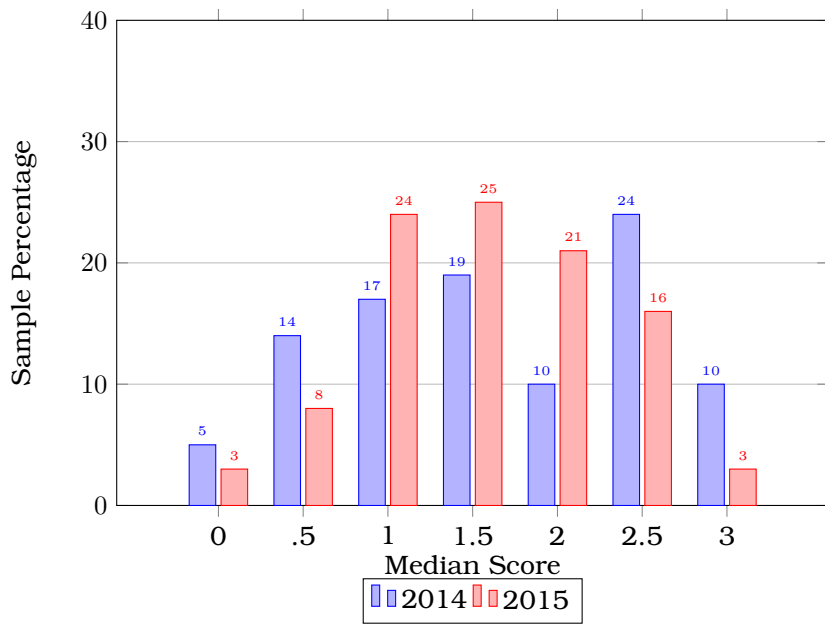


Figure 6: Sources and Evidence

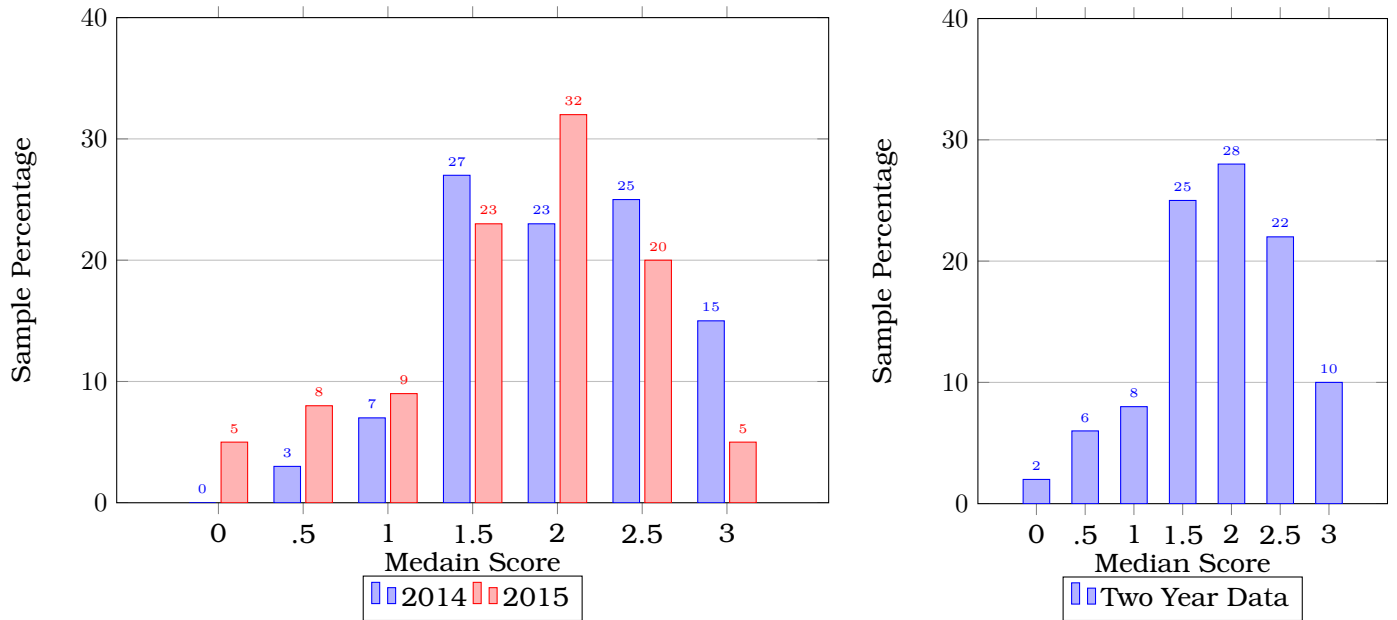
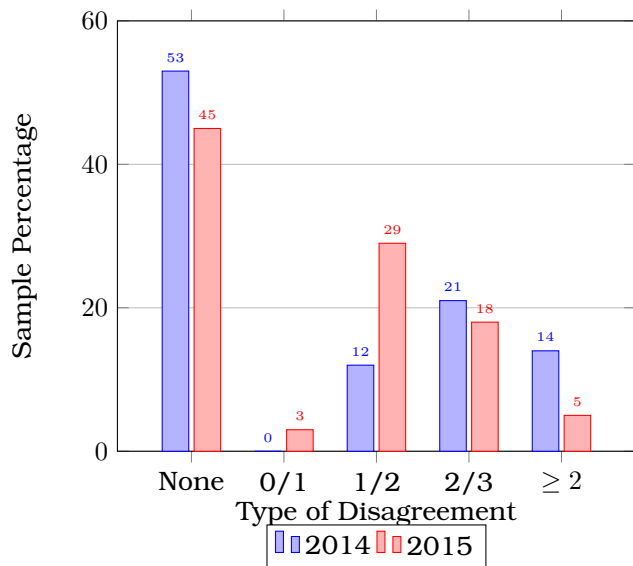


Figure 7: Control of Syntax and Mechanics

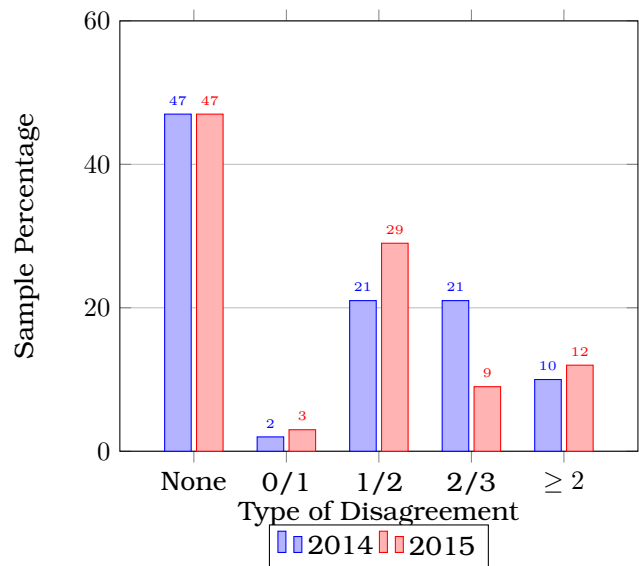
### 2.2.2 Rater Agreement

The bar charts below give the type of disagreement for each rubric item. Perfect agreement is categorized as None; the three types of adjacent agreement are categorized as 0/1, 1/2, and 2/3; while disagreement by two or more units is categorized as  $\geq 2$ .

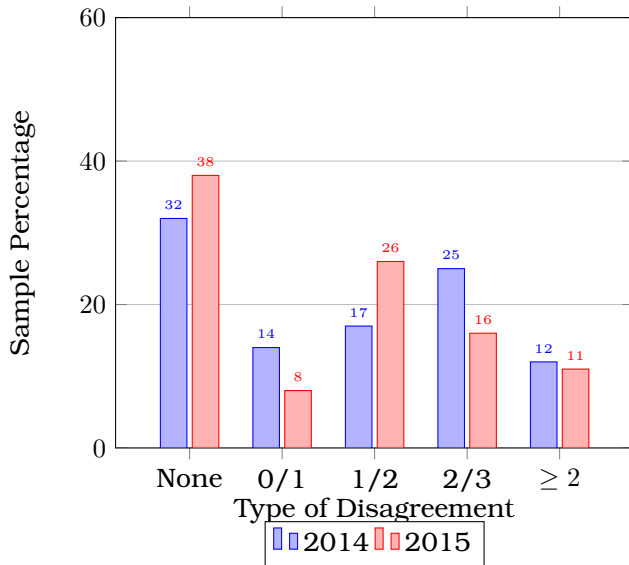
In all cases, the data suggests that if raters assigned adjacent scores they were more likely to fall in either the 1/2 or 2/3 categories than the 0/1 category. Also worth noting is the relatively high percentage of 2015 artifacts on which raters disagreed by two or more units on the Content Development and Sources and Evidence rubric items.



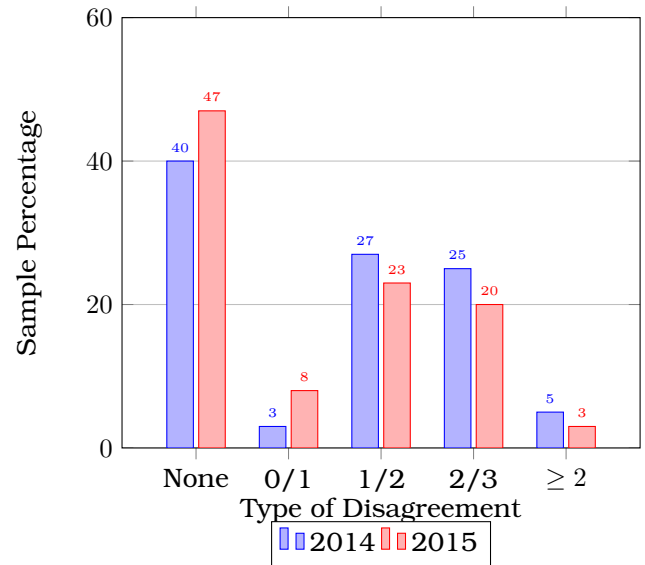
(a) Purpose for Writing



(b) Content Development



(c) Sources and Evidence



(d) Control of Syntax and Mechanics

Figure 7: Type of Disagreement - Written Communication

### 3 Critical and Creative Thinking

#### 3.1 Summary of the Critical and Creative Thinking Outcome

The data gathered in 2014 and 2015 appears to suggest that our students are performing best on the Formulates a Significant Question, Problem or Issue and Demonstrates Understanding of the Problem, Question or Issue rubric items as evidenced by a high percentage of high median scores (median scores between 2 and 3 inclusive) and a low percentage of low median scores (median scores of 0 or 0.5). Also worth noting are the rubric items on which our students are performing poorly, namely the Considers Underlying Assumptions; Considers Alternate, Potentially Divergent or Contradictory Perspectives; and Produces Something Original rubric items, as evidenced by the low percentage of high median scores and high percentage of low median scores. However, in all three of these areas students performed better in 2015 than in 2014.

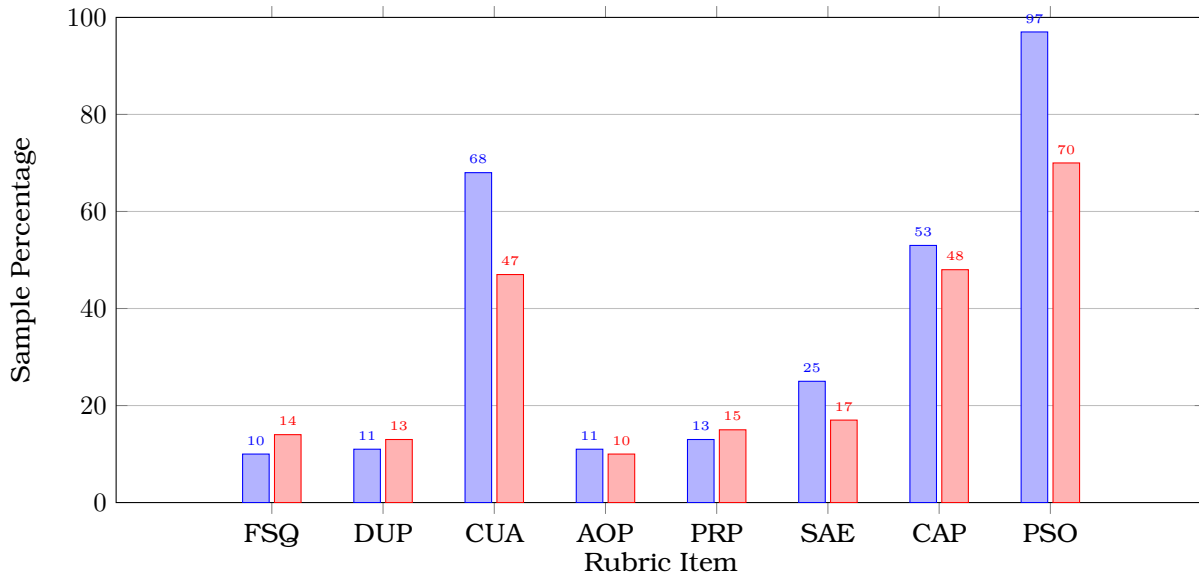


Figure 8: Sample Proportions - Critical and Creative thinking,  $0 \leq \text{Median score} \leq 0.5$

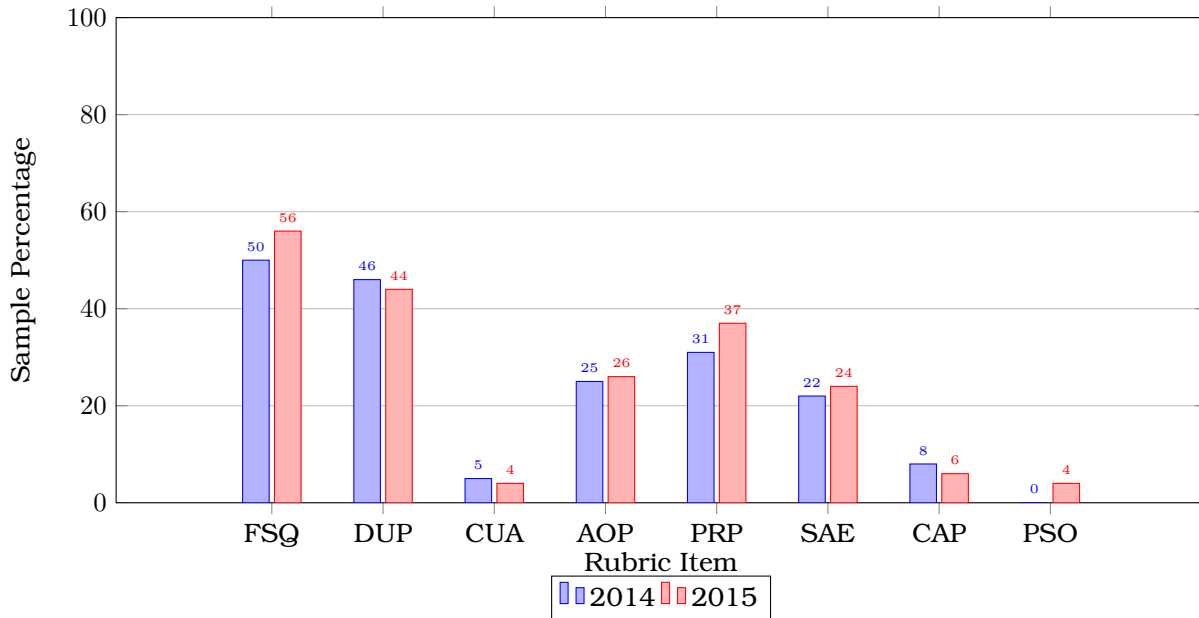


Figure 9: Sample Proportions - Critical and Creative thinking,  $2 \leq \text{Median score} \leq 3$

In 2015, five of the eight Critical and Creative Thinking rubric items (Formulates a Significant Question, Problem or Issue; Demonstrates Understanding of the Problem, Question or Issue; Provides Reason for Position; Selects and then Analyzes Evidence for Reasoning; and Produces Something Original) exhibited good agreement ( $AC2 > 0.60$ ). On the other hand, two of the three rubric items that had good agreement in 2014, dropped to moderate agreement in 2015 (Considers Underlying Assumptions and Considers Alternate, Potentially Divergent or Contradictory Perspectives).

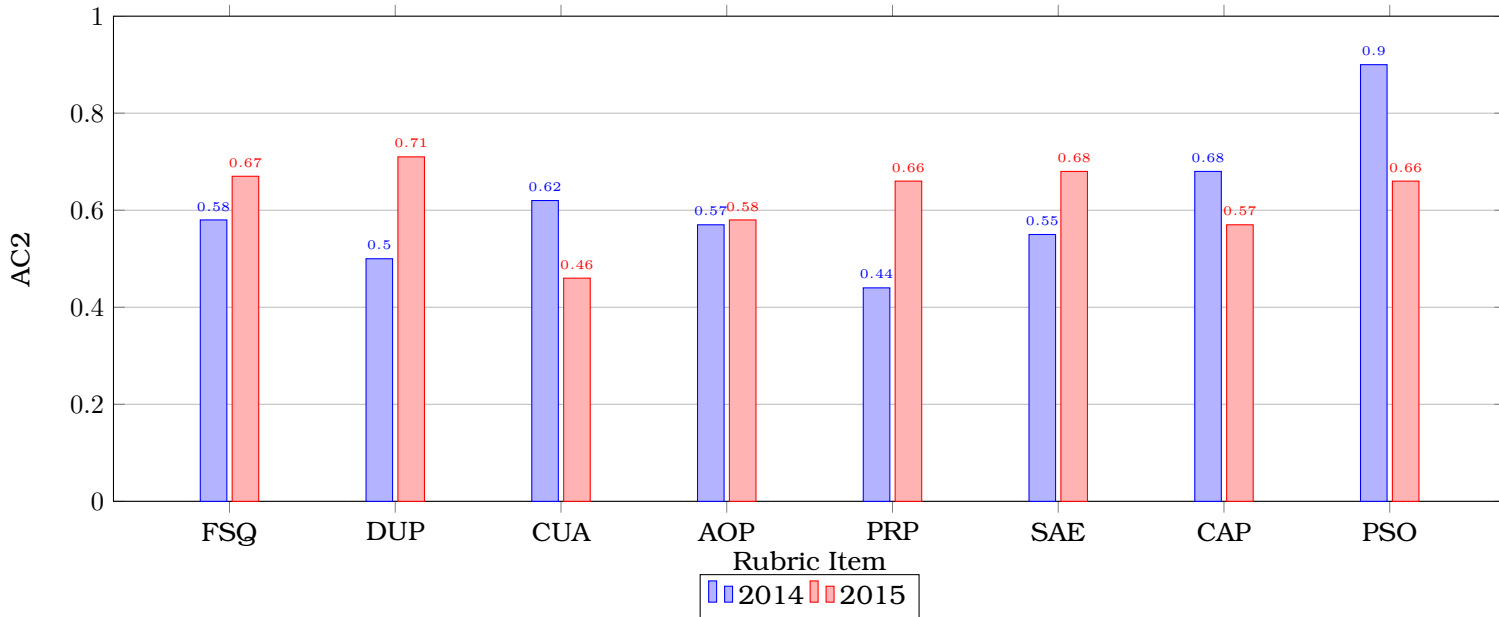


Figure 10: Gwet's AC2 - Critical and Creative Thinking

The 95% confidence intervals below use the 2015 sample data to give an interval estimate for the median scores of all RIC freshmen on the Critical and Creative Thinking rubric items.

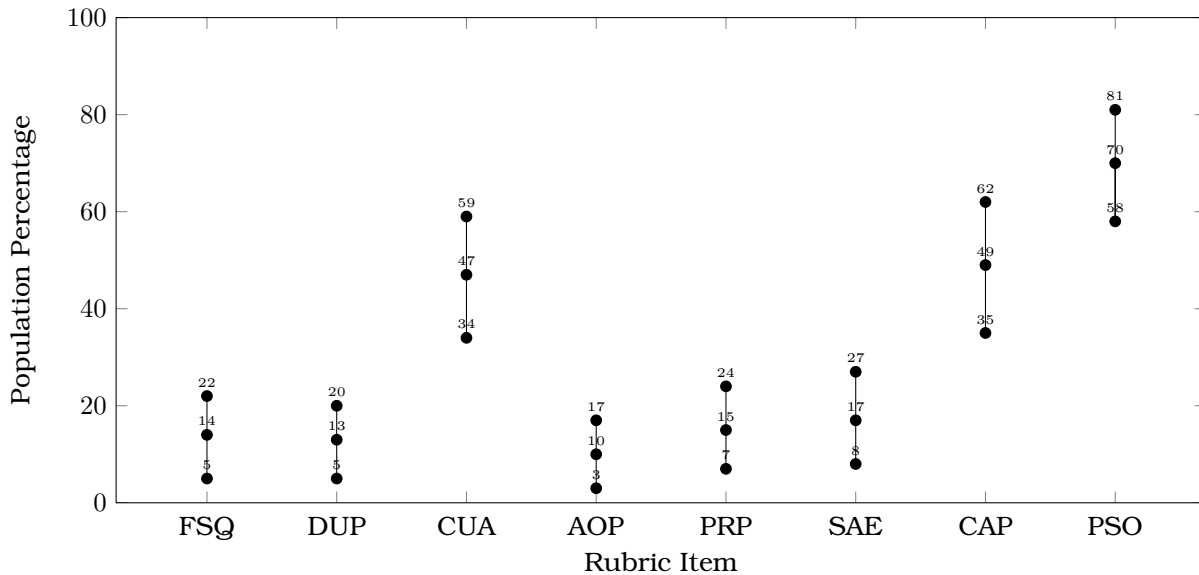


Figure 11:  $0 \leq \text{Median score} \leq 0.5$

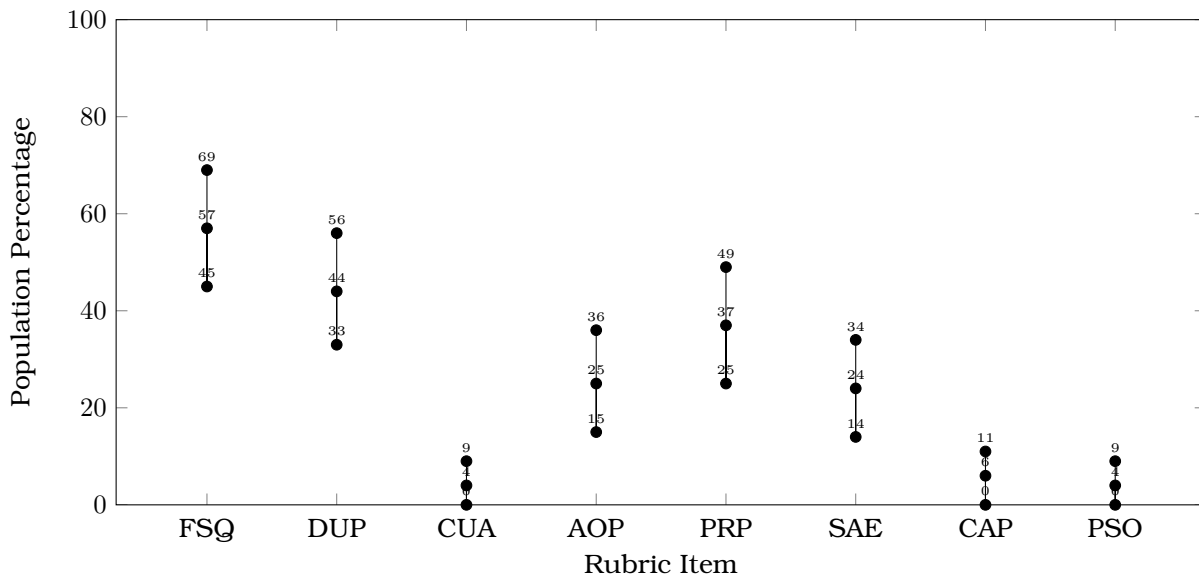


Figure 12:  $2 \leq \text{Median score} \leq 3$

### 3.2 Critical and Creative Thinking - Summaries by Rubric Item

#### 3.2.1 Student Performance

Once again the individual rubric data suggests that our students are not performing well on the Considers Underlying Assumptions; Considers Alternate, Potentially Divergent or Contradictory Perspectives; and Produces Something Original rubric items. For all three of these items the distribution of scores in the two year data charts is skewed to the right indicating more low than high median scores on these items.



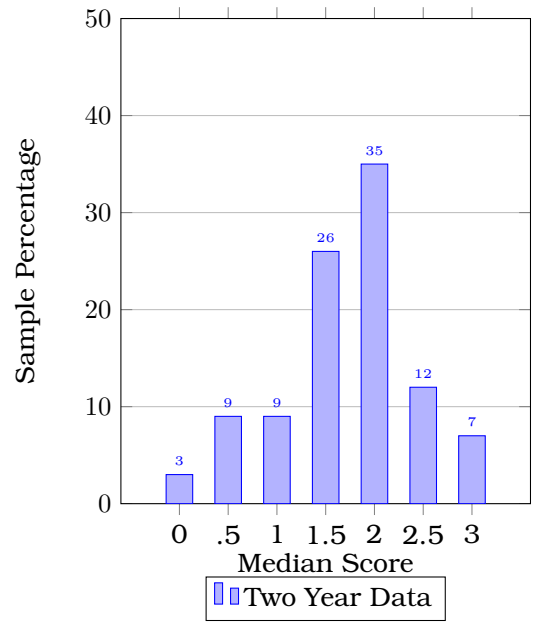
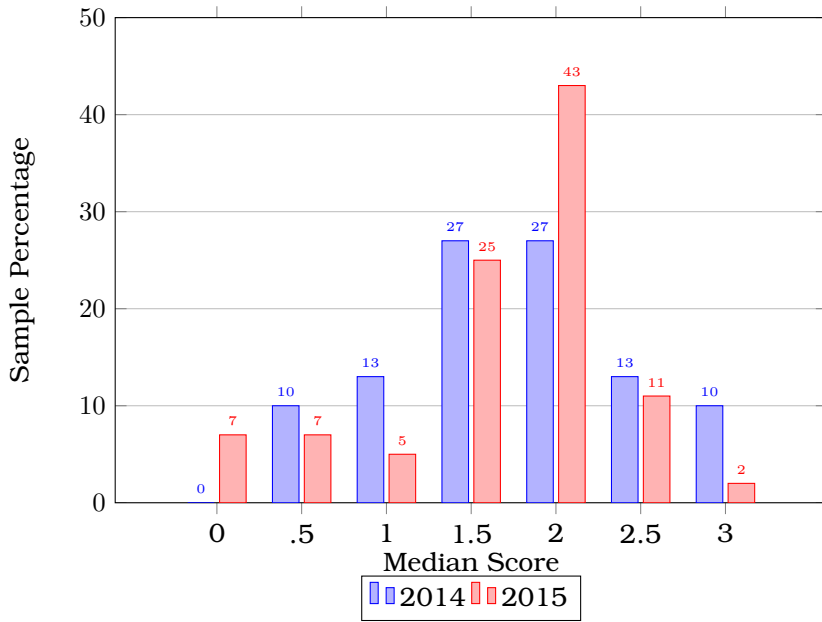


Figure 13: Formulates a Significant Question, Problem or Issue

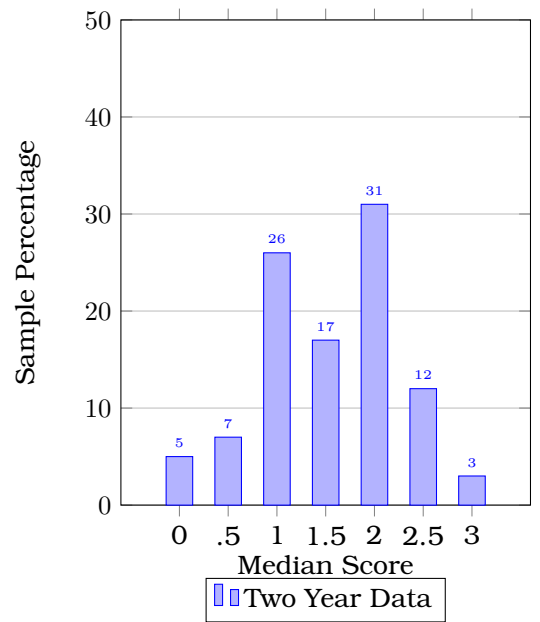
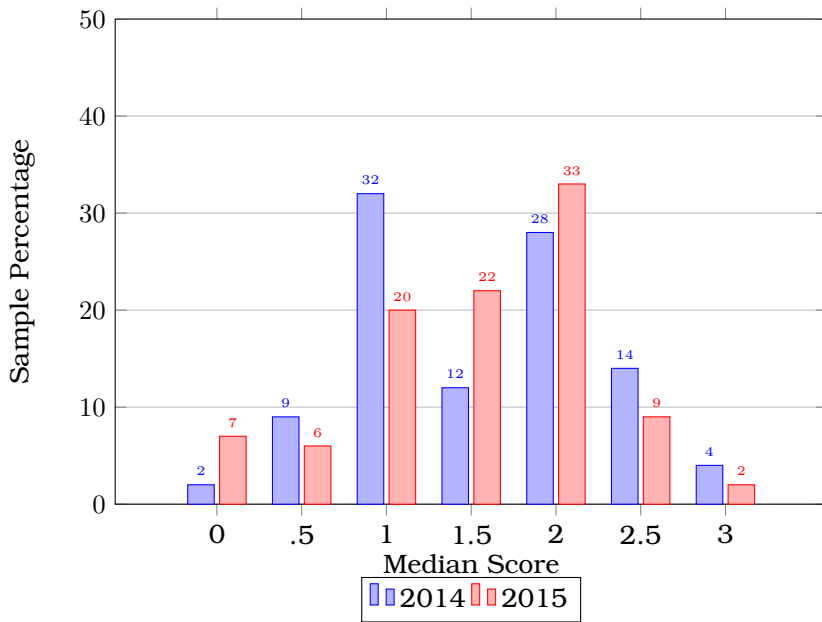


Figure 14: Demonstrates Understanding of the Problem, Question or Issue

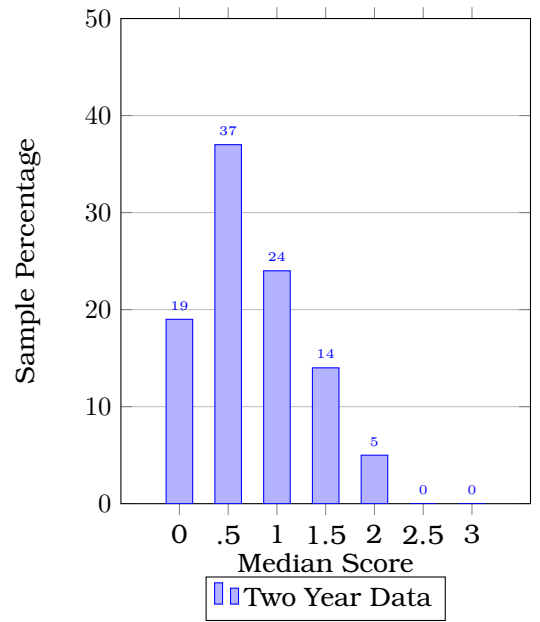
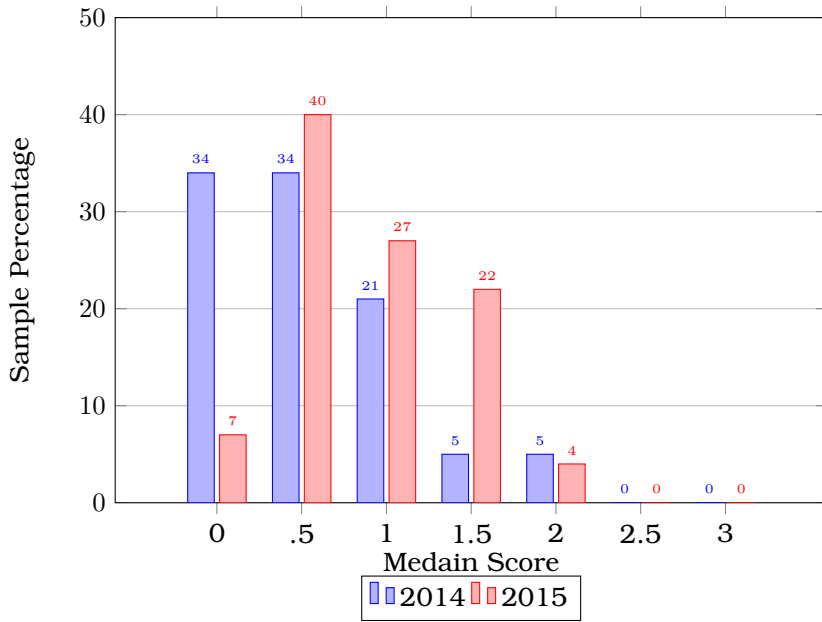


Figure 15: Considers Underlying Assumptions

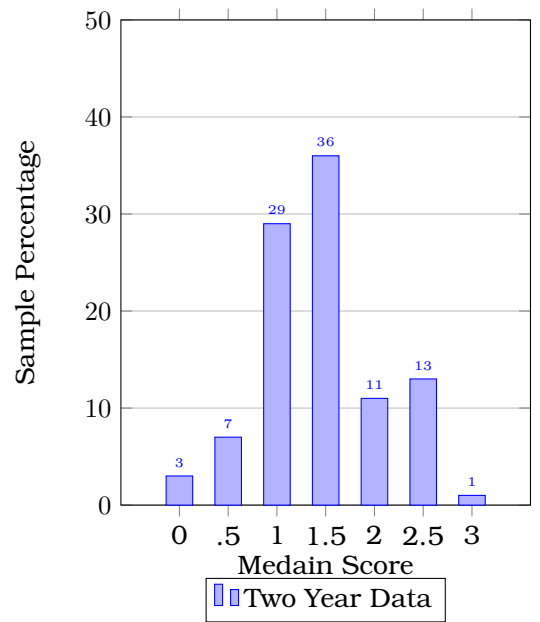
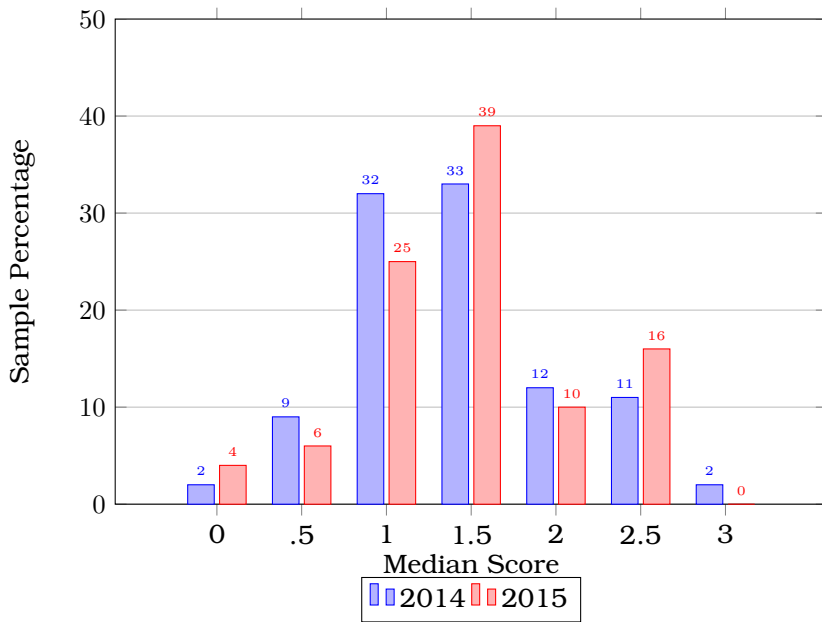


Figure 16: Articulates Own Perspective, Hypothesis or Position

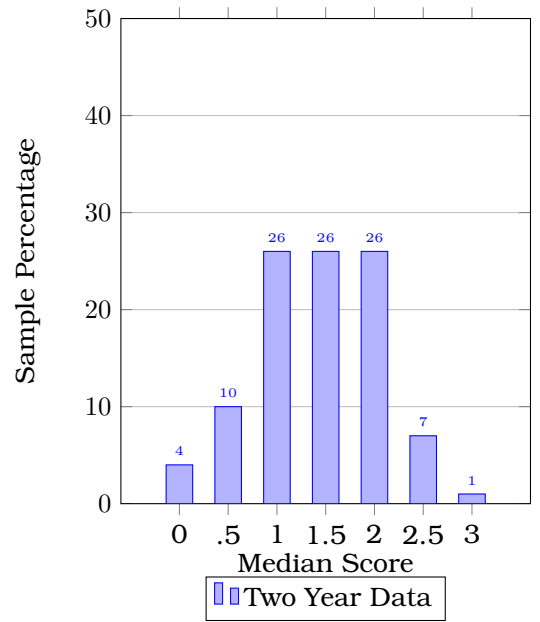
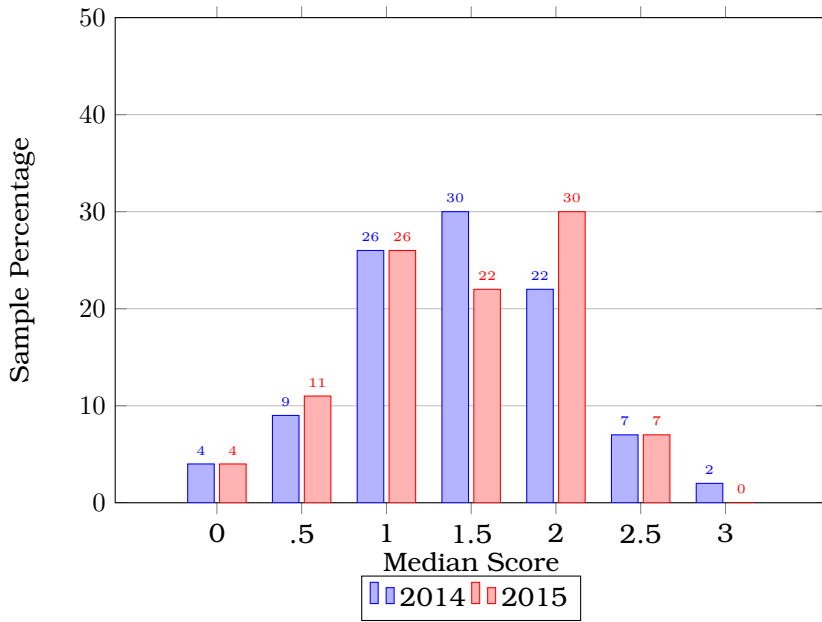


Figure 17: Provides Reasons for Position

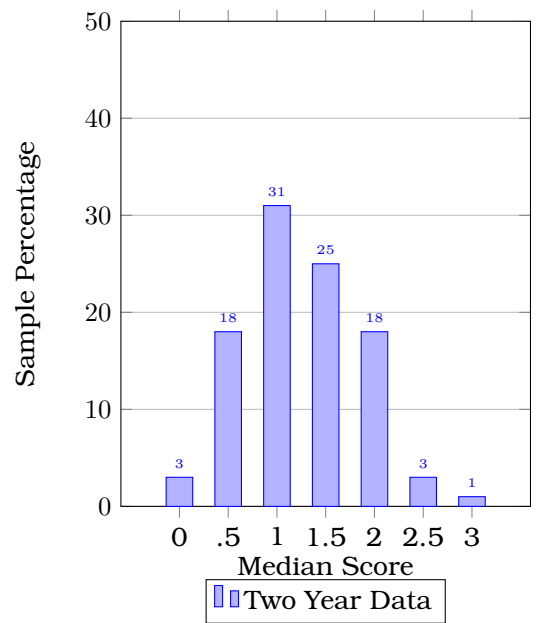
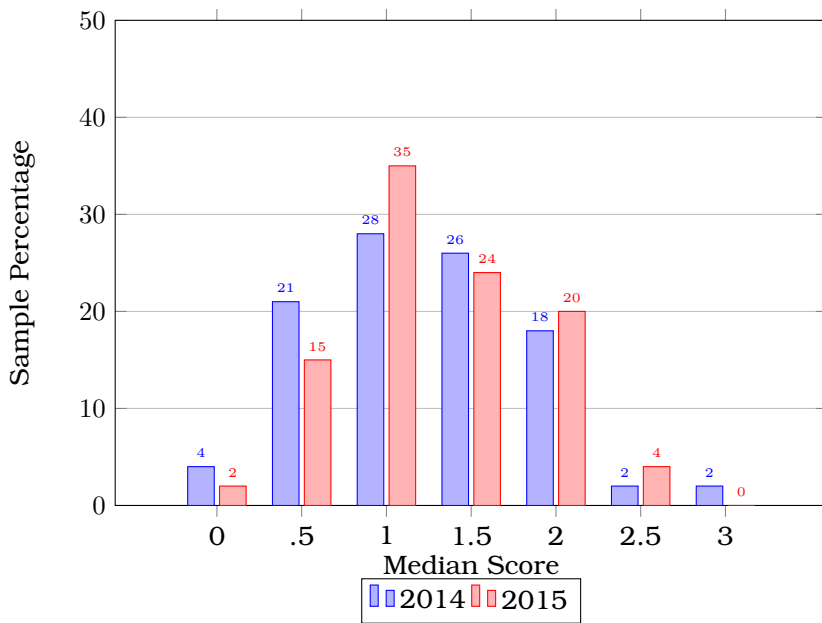


Figure 18: Selects and then Analyzes Evidence for Reasoning

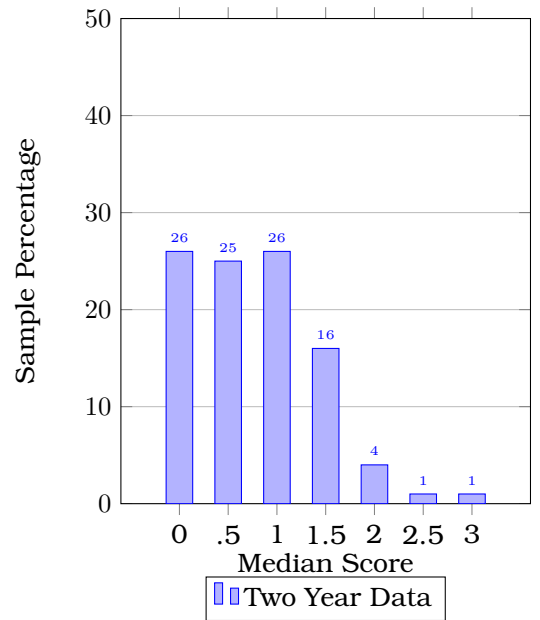
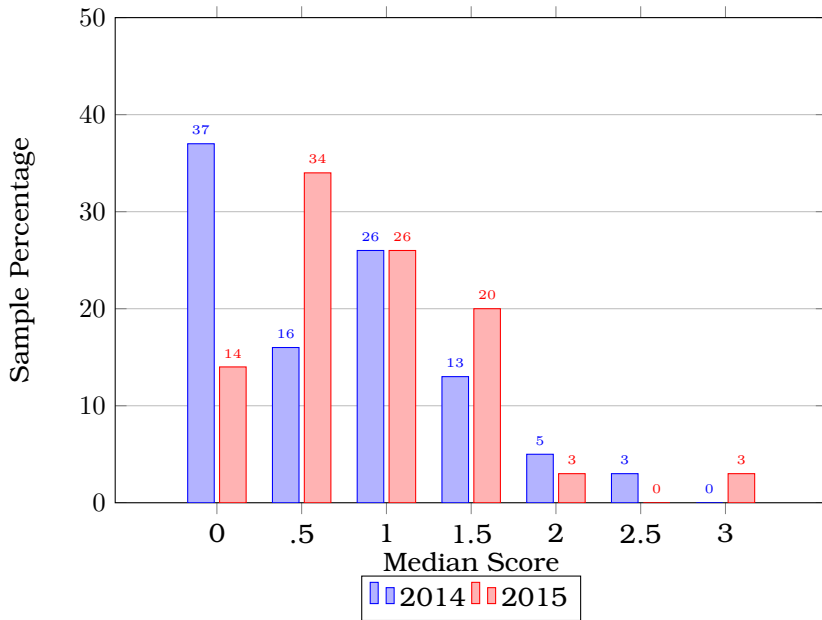


Figure 19: Considers Alternate, Potentially Divergent or Contradictory Perspectives

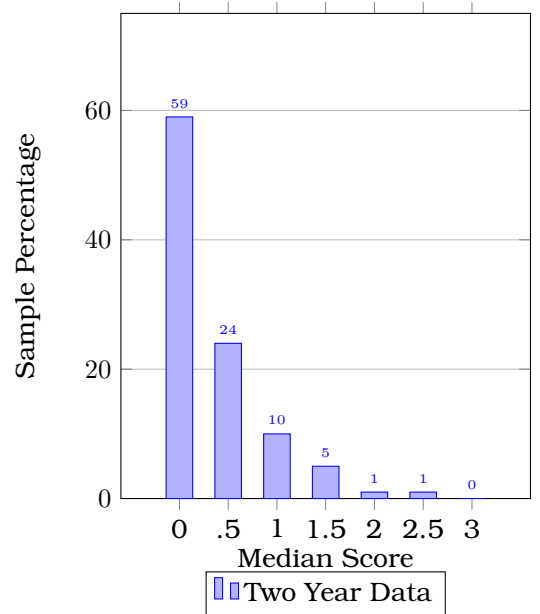
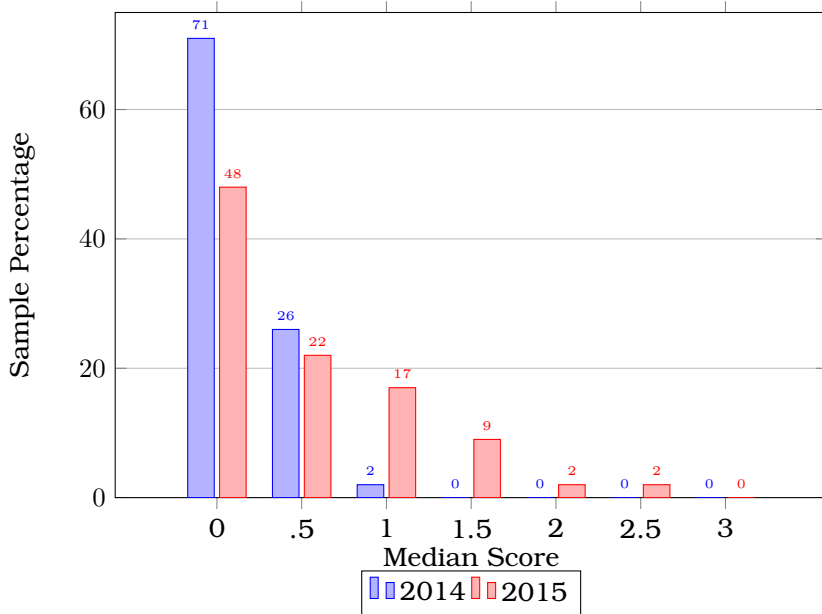
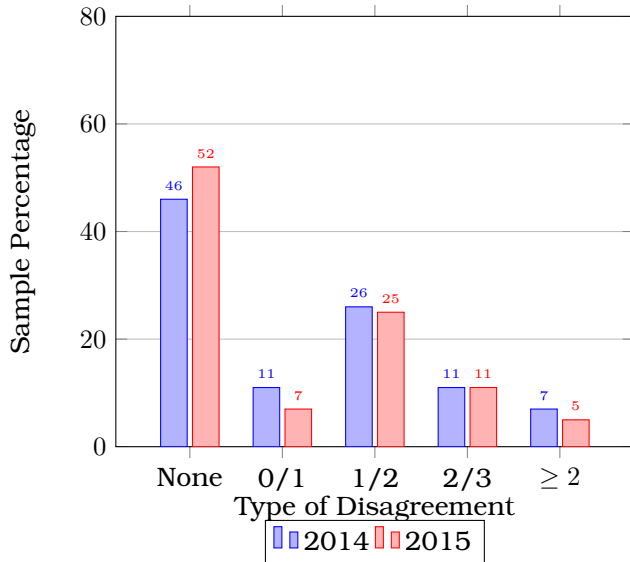


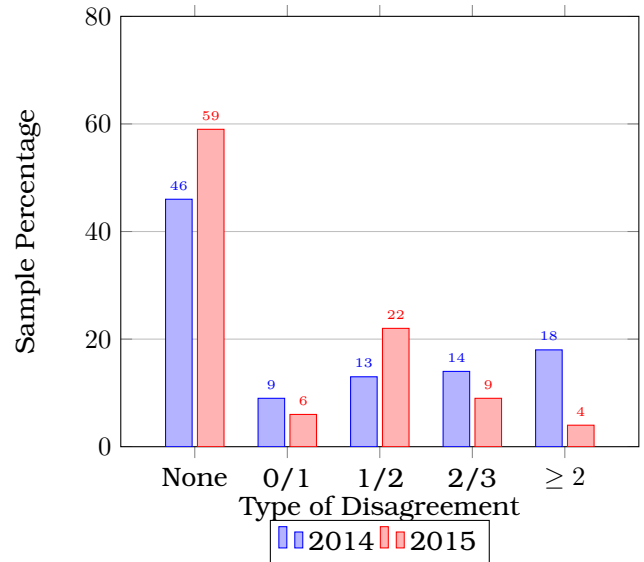
Figure 20: Produces Something Original

### 3.2.2 Rater Agreement

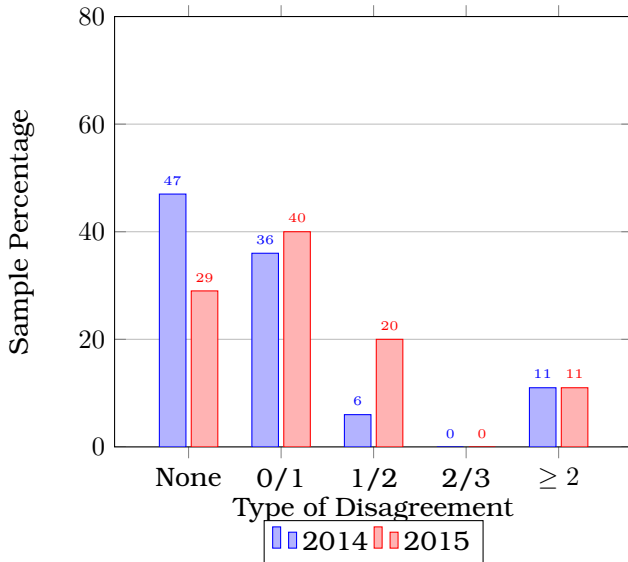
A good deal can be learned about our rubrics not only by looking at percent agreement, but by focusing on disagreement, specifically the exact ways in which raters disagree. For example, while for most rubric items a relatively low percentage of raters assigned adjacent scores of 0 and 1, for the Considers Underlying Assumptions rubric item this percentage seems higher than normal in both 2014 and 2015. This trend may indicate that we need to clarify the difference between a rubric score of 0 and of 1 for this particular item. Also noticeable is the high percentage of artifacts on which raters assigned adjacent scores of 1 and 2 on the Articulates Own Perspective, Hypothesis or Position rubric item in both 2014 and 2015.



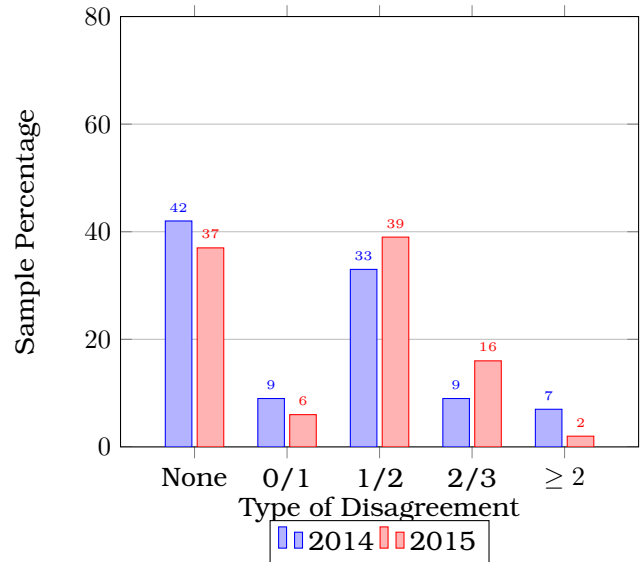
(a) Formulates a Significant Question, Problem or Issue



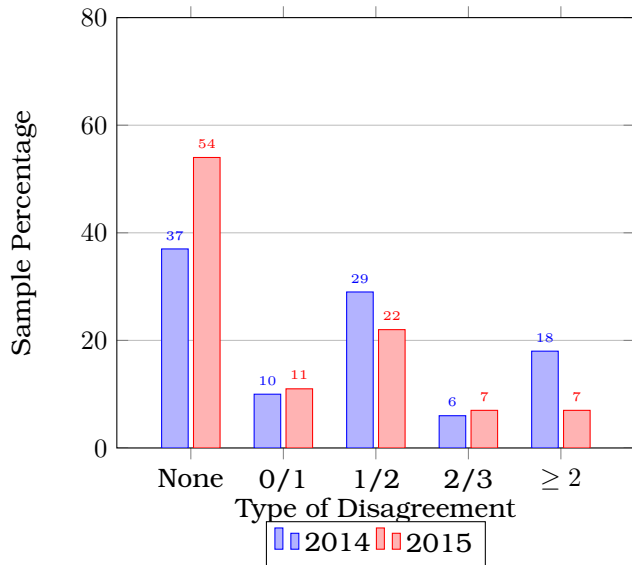
(b) Demonstrates Understanding of the Problem, Question or Issue



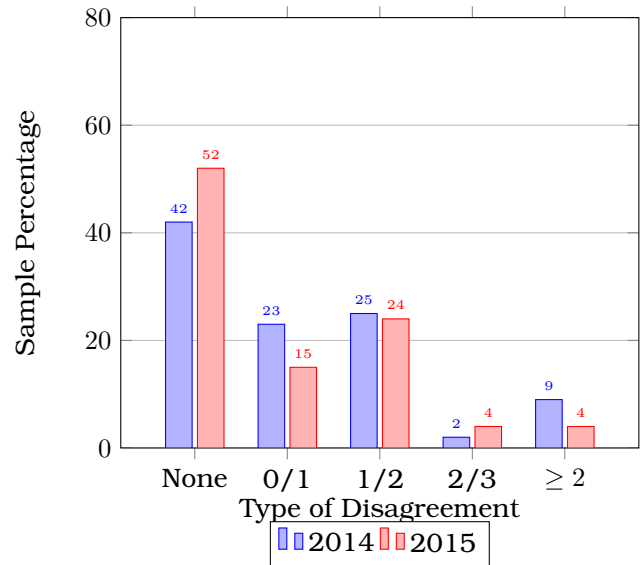
(c) Considers Underlying Assumptions



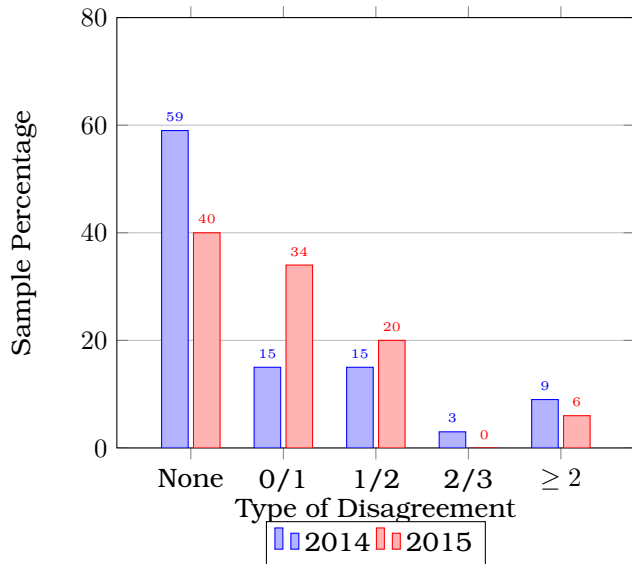
(d) Articulates Own Perspective, Hypothesis or Position



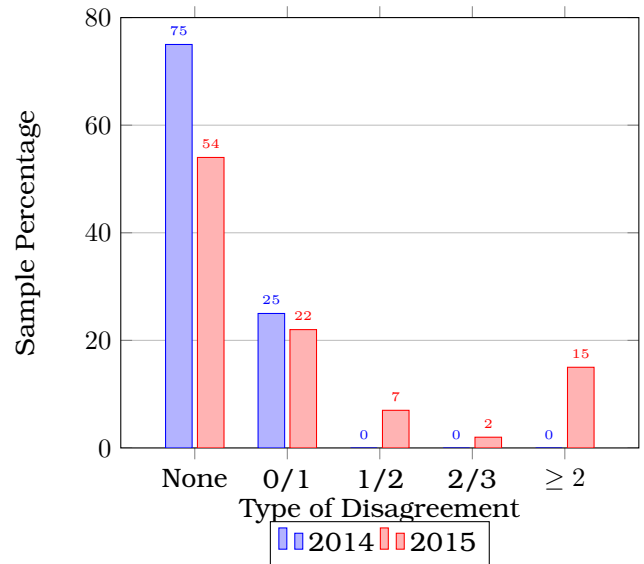
(e) Provides Reason for Position



(f) Selects and then Analyzes Evidence for Reasoning



(g) Considers Alternate, Potentially Divergent or Contradictory Perspectives



(h) Produces Something Original

Figure 20: Type of Disagreement

## 4 Research Fluency

### 4.1 Summary of the Research Fluency Outcome

The data suggests that in 2015 our students did not perform as well on the rubric item 3: Evaluates all information critically, including its sources and authority, parts b: Recognizes points of view in or quality of material and c: Responds to points of view in or quality of material as they did on the other Research Fluency rubric items.

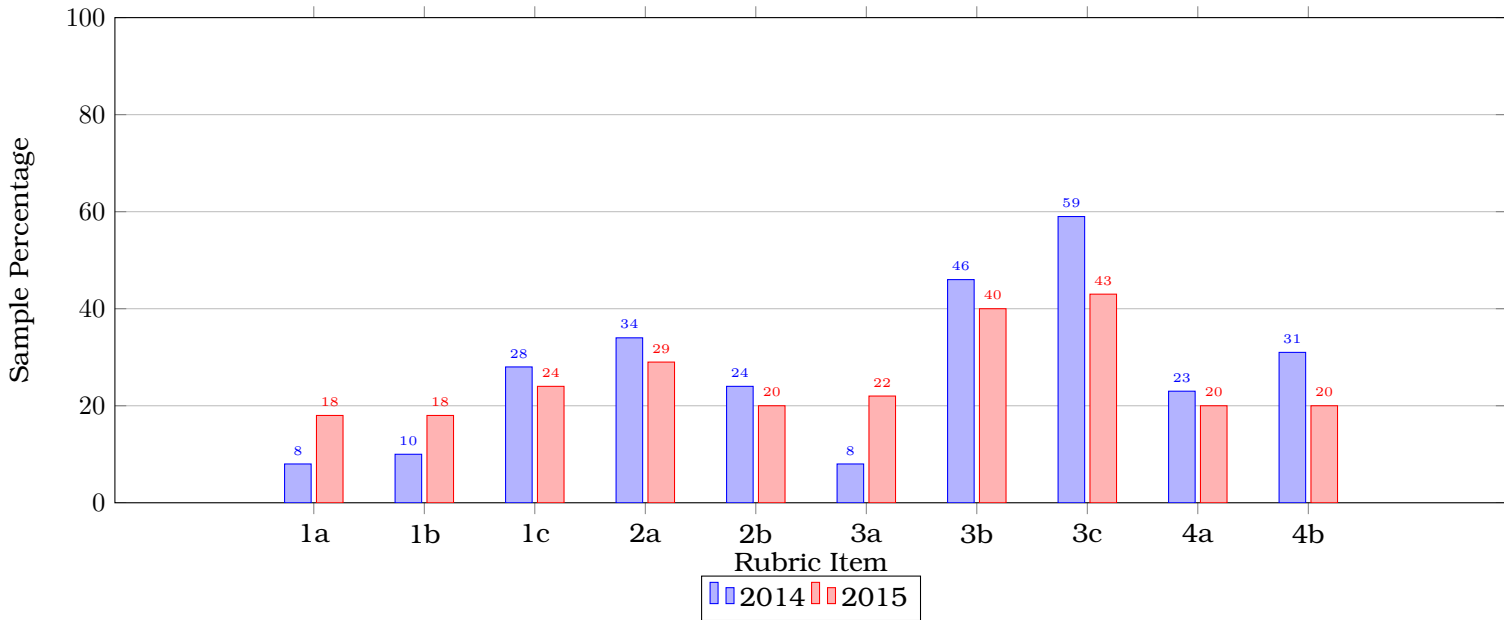


Figure 21: Sample Proportions - Research Fluency  $0 \leq \text{Median score} \leq 0.5$

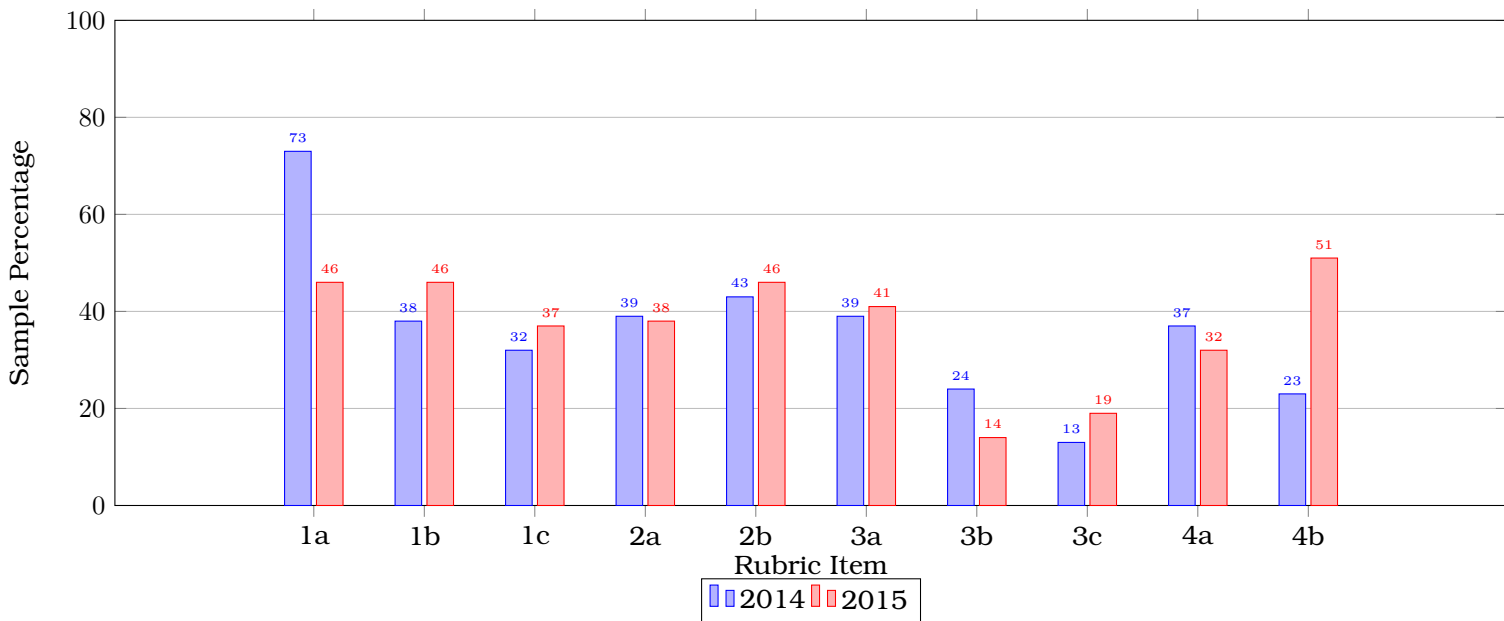


Figure 22: Sample Proportions - Research Fluency  $2 \leq \text{Median Score} \leq 3$

In 2015, the data shows evidence of an increase in agreement for seven of the ten research fluency rubric items. Two of the seven items (1a and 4a) increased from fair to moderate agreement, another two of the seven items (1b and 3a) increased from moderate to good agreement, while the remaining three items (1c, 2b and 3c) showed increase, but remained at the moderate level of agreement. Of the three items that showed a decrease in agreement item 4b dropped only slightly, remaining at the moderate level of agreement, while items 2a and 3b dropped from moderate to fair agreement.

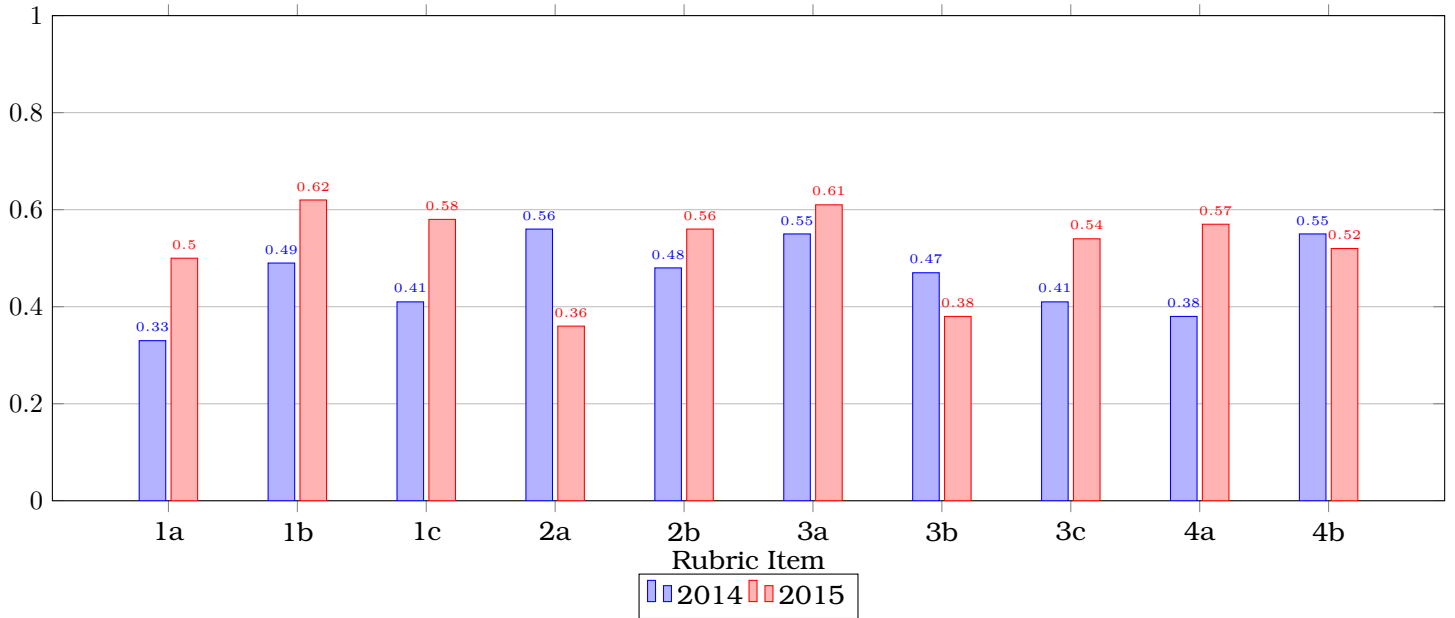


Figure 23: Gwet's AC2 - Research Fluency

The 95% confidence intervals below use the 2015 sample data to give an interval estimate for the median scores of all RIC freshmen on the Research Fluency rubric items.

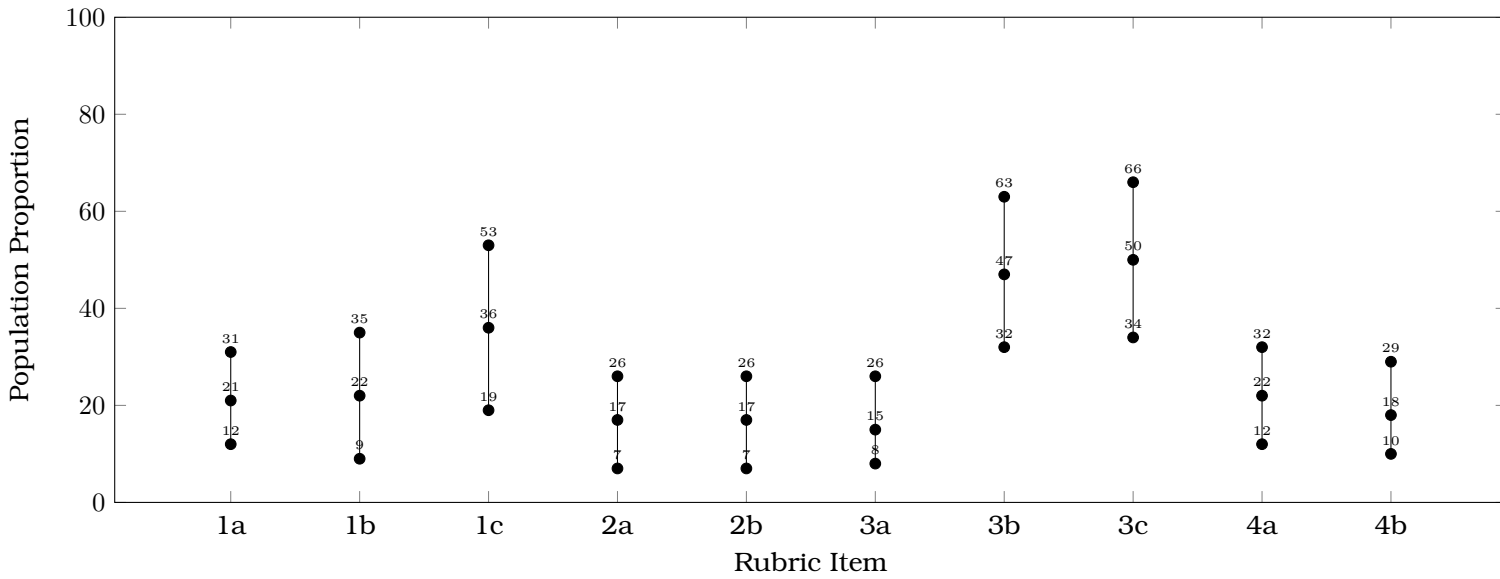


Figure 24: Confidence Intervals - Research Fluency  $0 \leq \text{Median score} \leq .5$



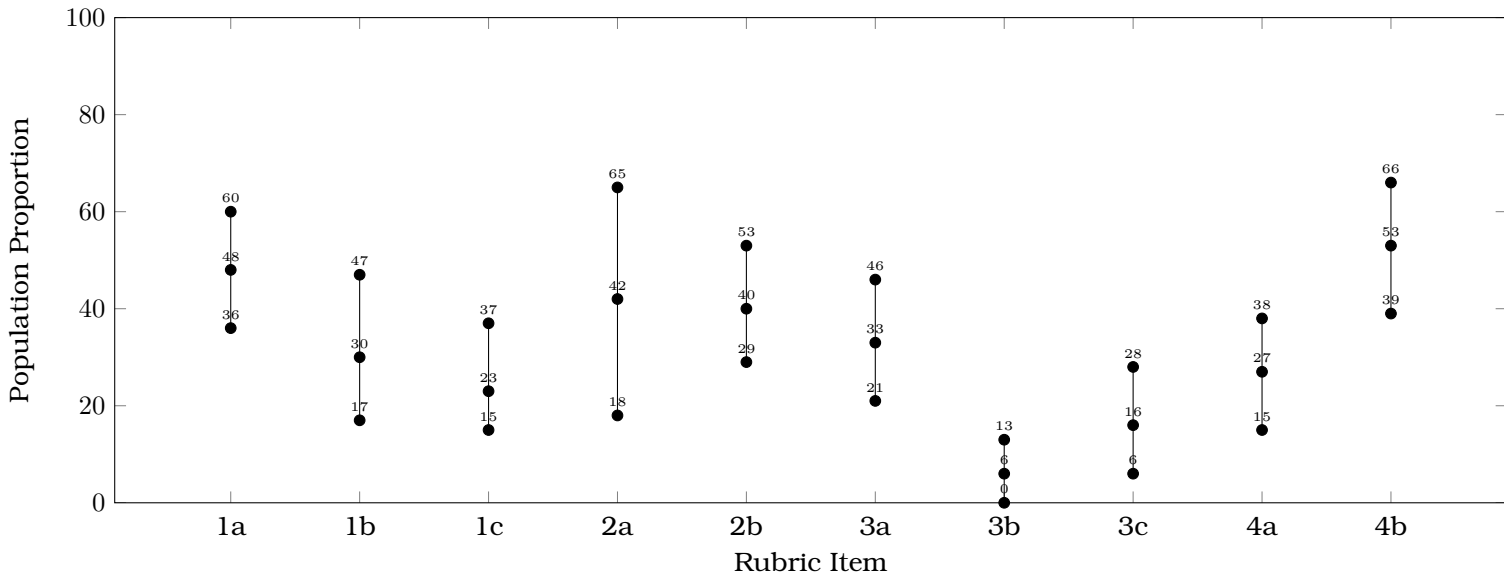


Figure 25: Confidence Intervals - Research Fluency  $2 \leq \text{Median score} \leq 3$

## 4.2 Research Fluency - Summaries by Rubric Item

### 4.2.1 Student Performance

The combined two year data charts that follow for each rubric item indicate that the distribution of scores for items 3b and 3c is skewed to the right meaning these rubric items had more low than high median scores.

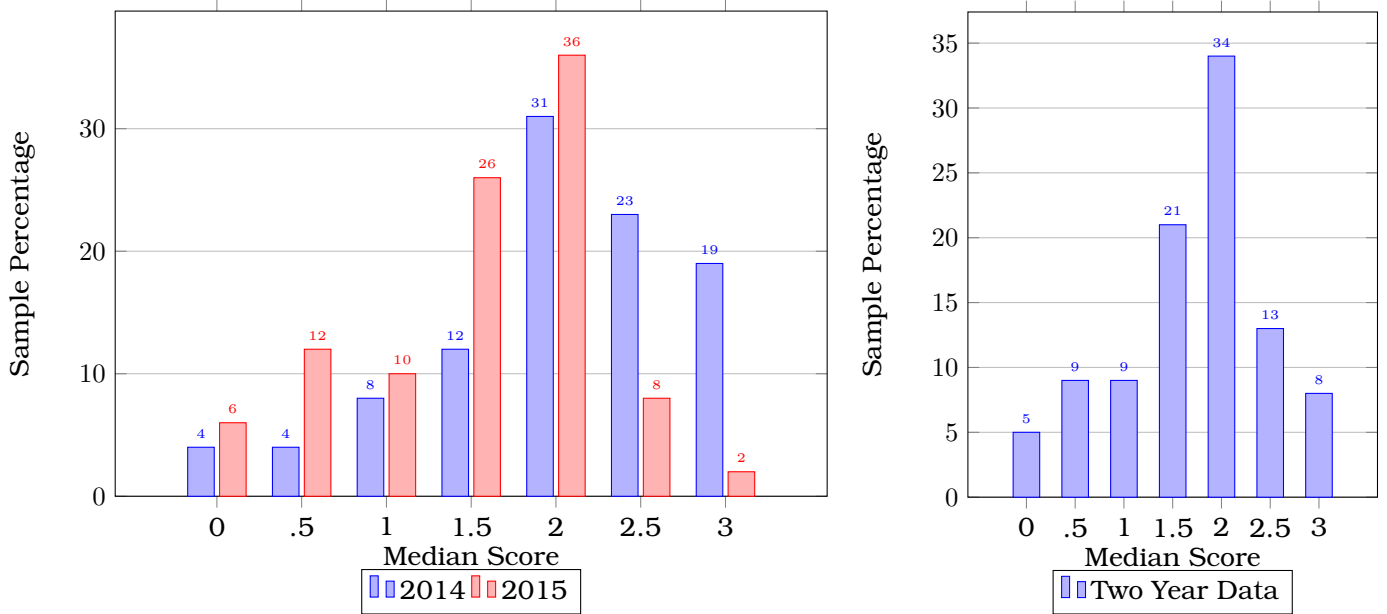


Figure 26: 1: Ability to Access Information, a: Define scope of research question thesis or information needed.

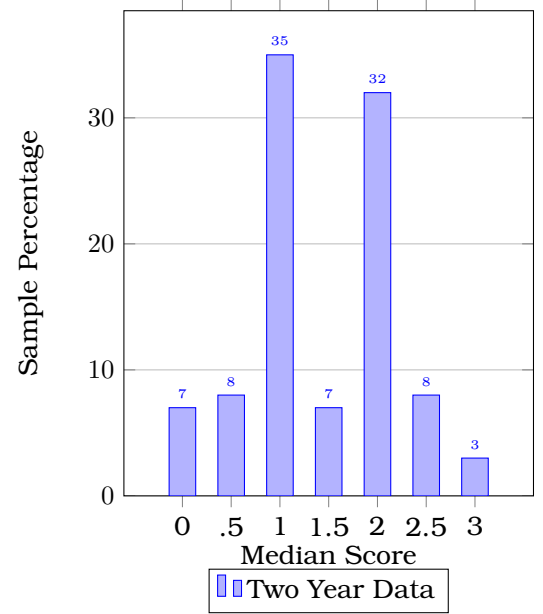
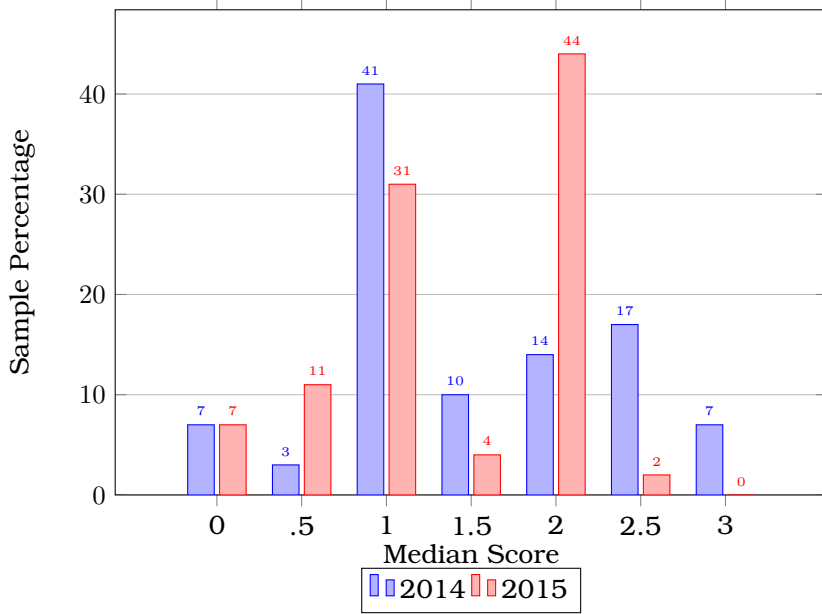


Figure 27: 1: Ability to Access Information, b: Effectively use tools appropriate for a specific task.

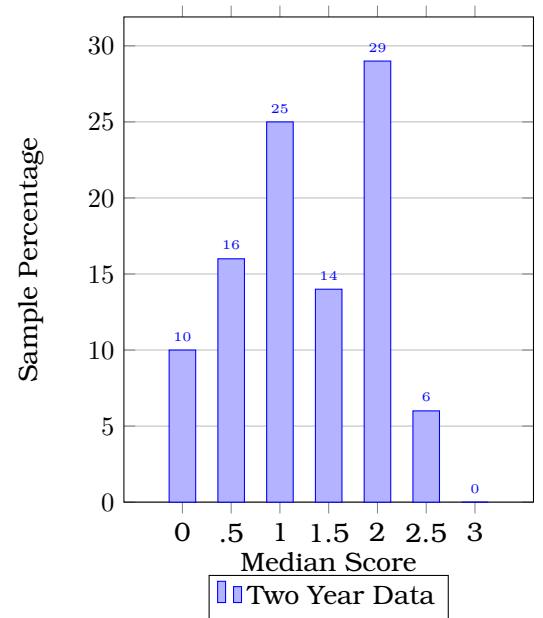
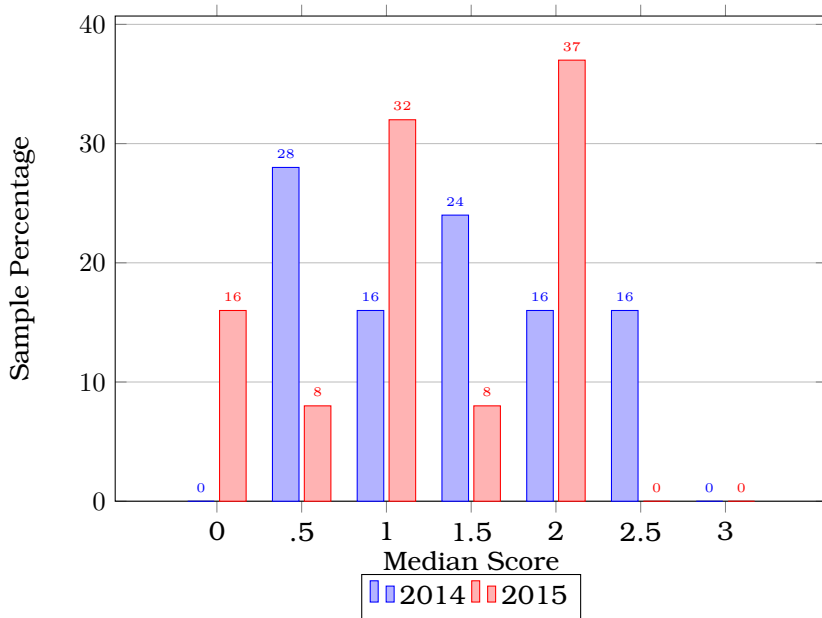


Figure 28: 1: Ability to Access Information, c: Identify key concepts for effective search strategy.

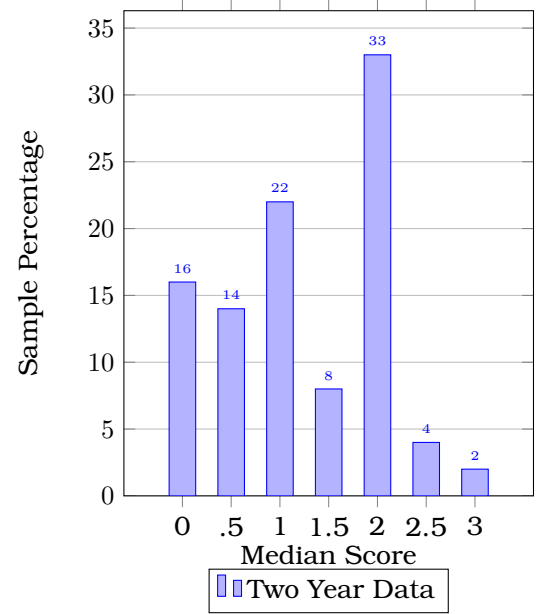
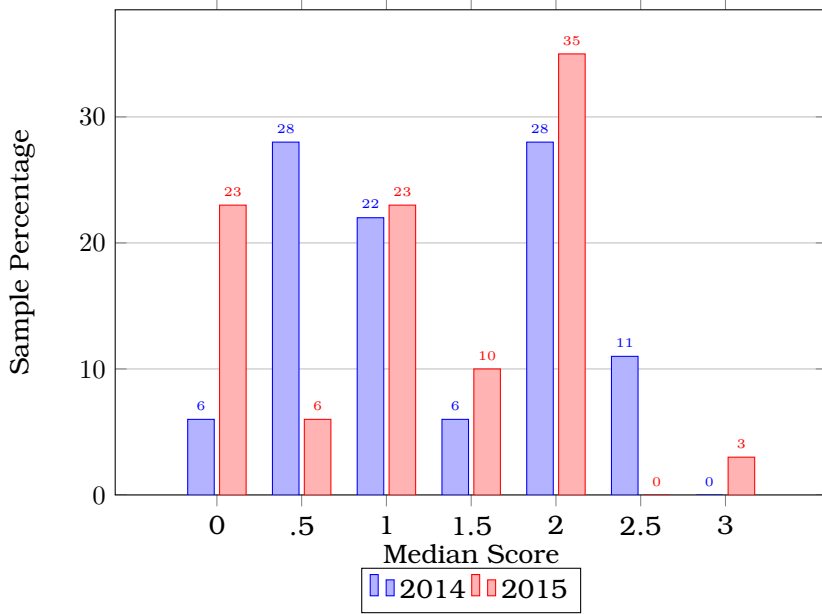


Figure 29: 2: Demonstrate understanding of sources used and information found, a: Differentiate among tertiary, secondary, and primary materials.

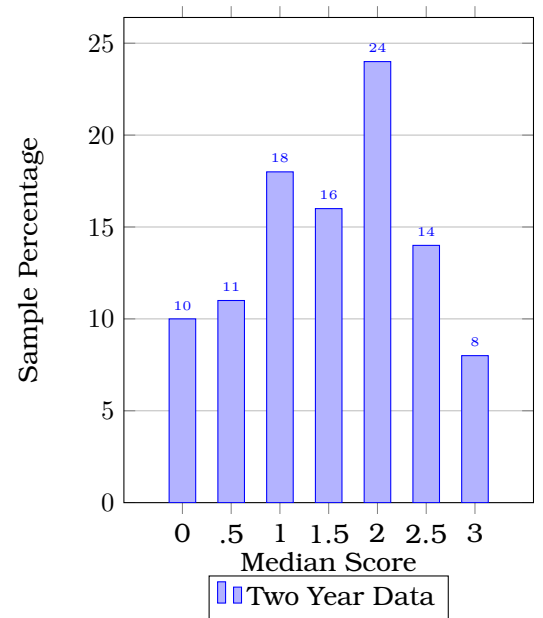
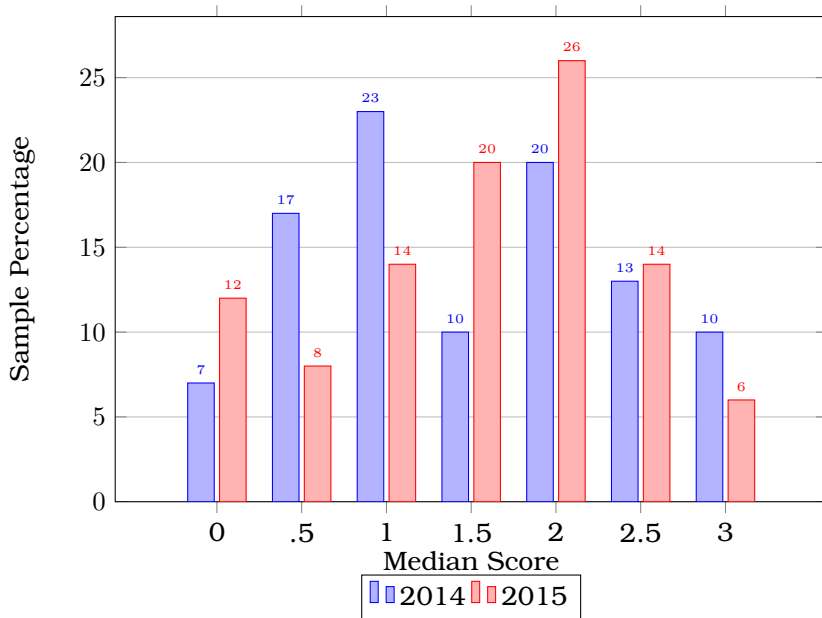


Figure 30: 2: Demonstrate understanding of sources used and information found, b: Select material or sources, appropriate to the task, considering appropriateness of popular/general sources vs scholarly/academic sources.

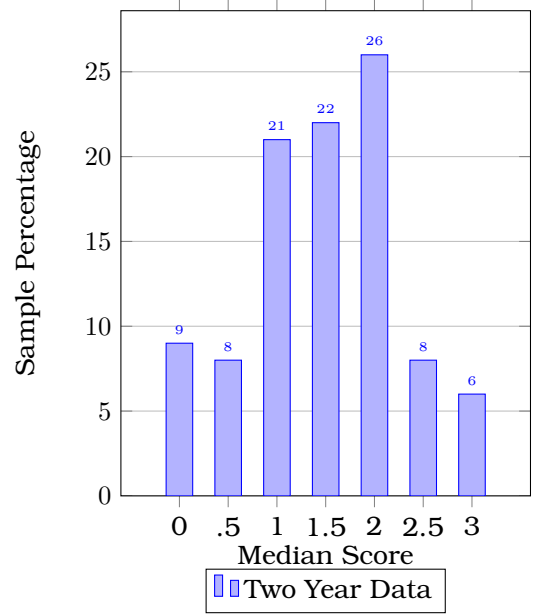
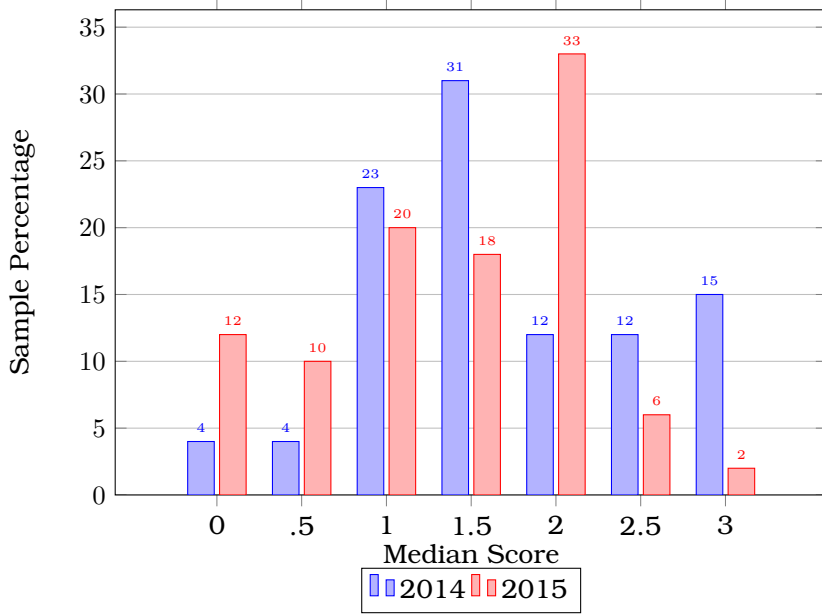


Figure 31: 3: Evaluate all information critically, including its sources and authority, a: Use relevant sources to address the research question, topic, or task.

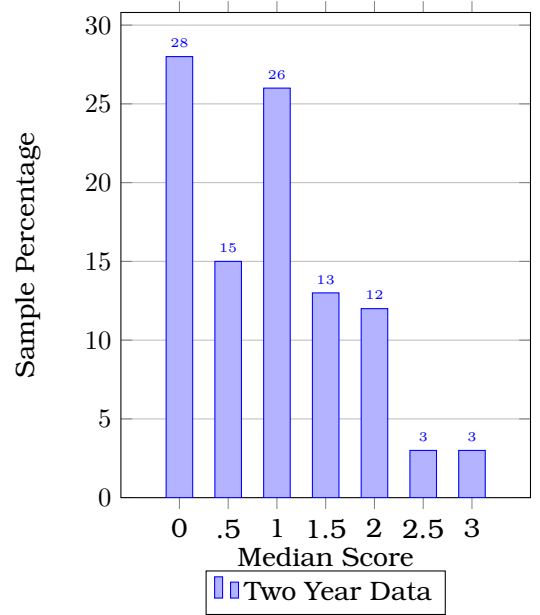
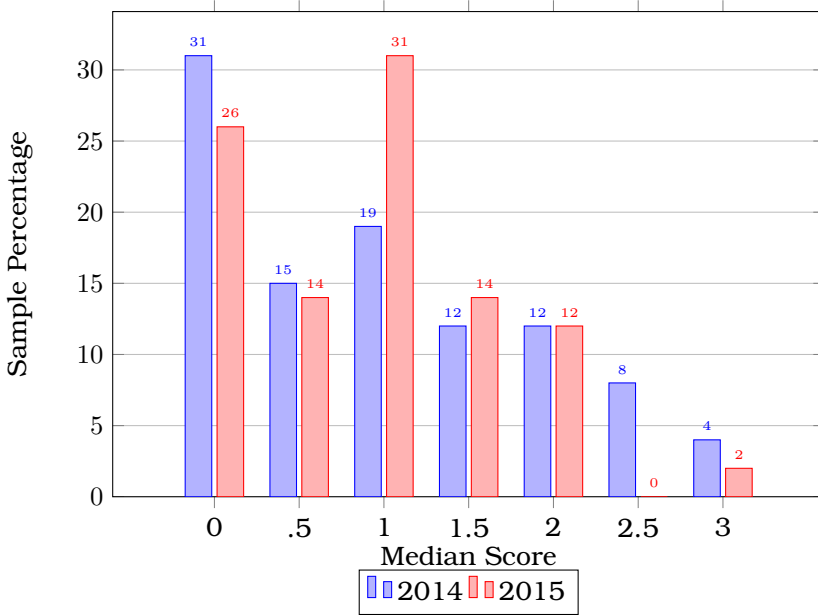


Figure 32: 3: Evaluate all information critically, including its sources and authority, b: Recognize point of view in or quality of material.

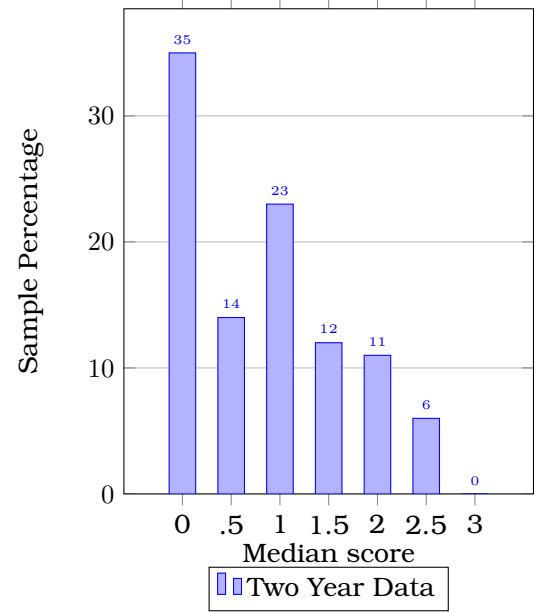
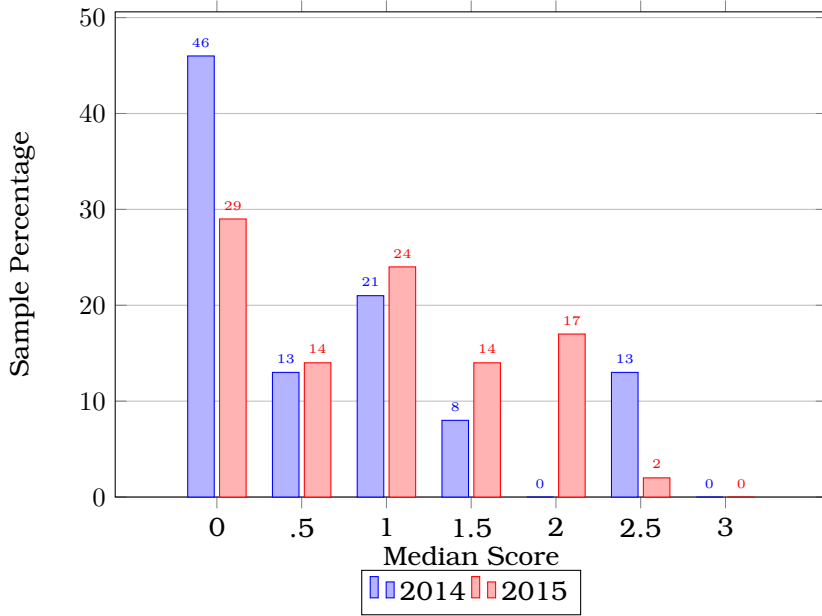


Figure 33: 3: Evaluate all information critically, including its sources and authority, c: Respond to point of view in or quality of material.

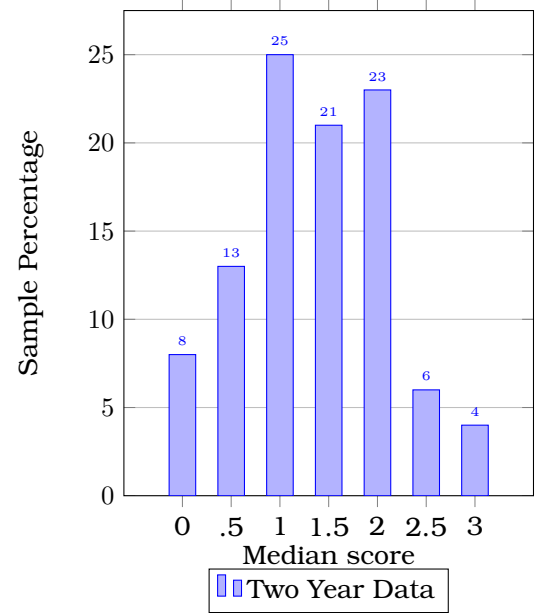
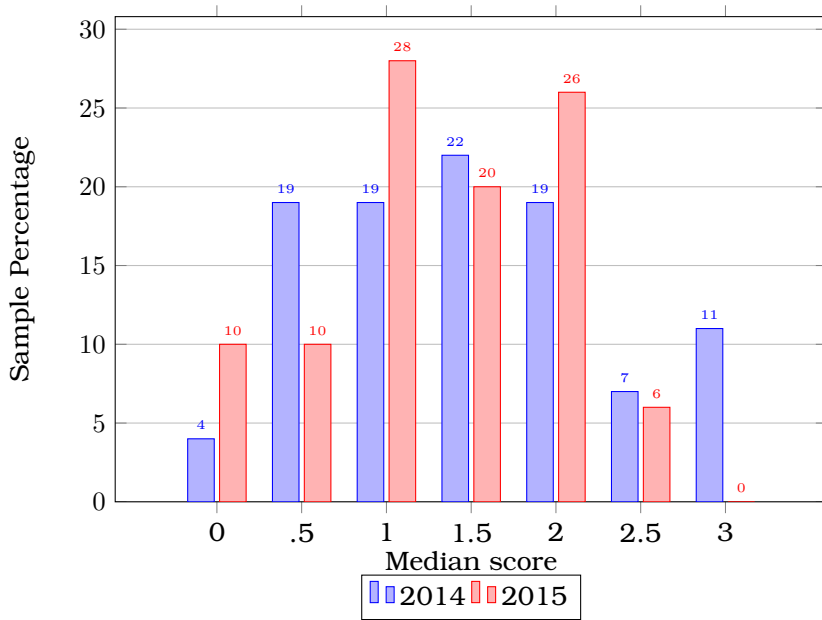


Figure 34: 4: Use information responsibly, a: Identify all sources using a consistent citation system.

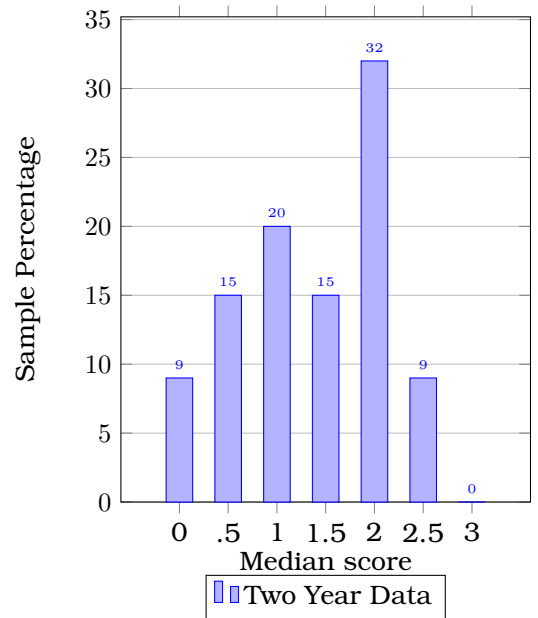
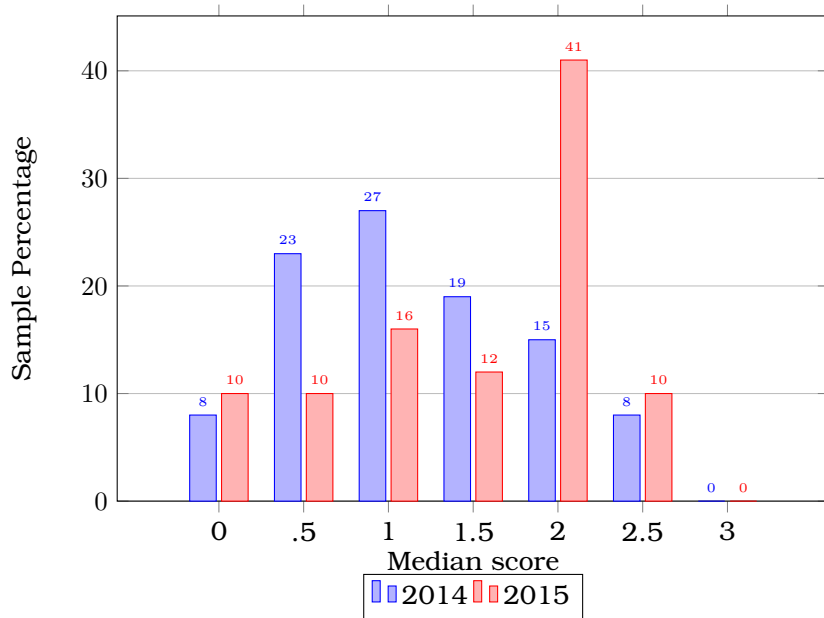
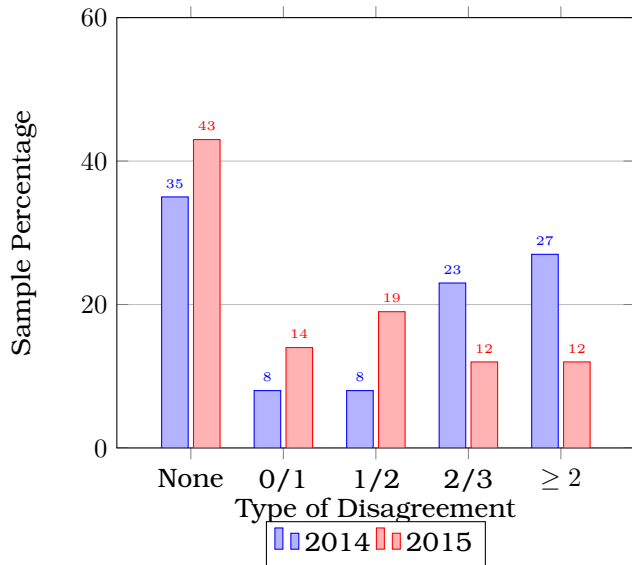


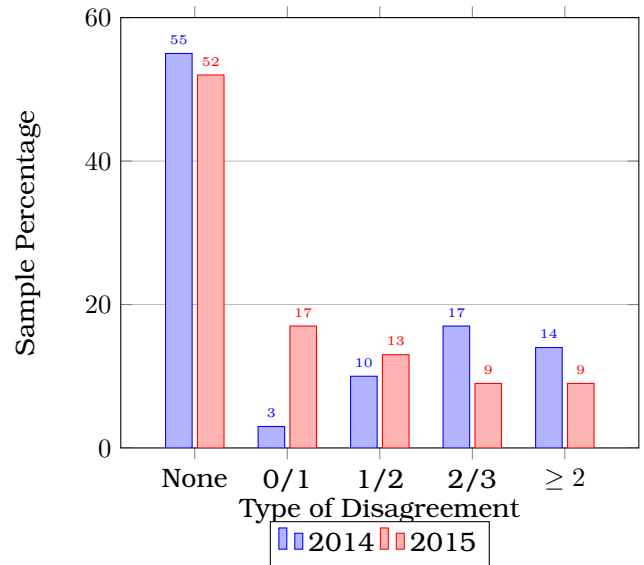
Figure 35: 4: Use information responsibly, b: Distinguish between common knowledge and ideas requiring attribution.

#### 4.2.2 Rater Agreement

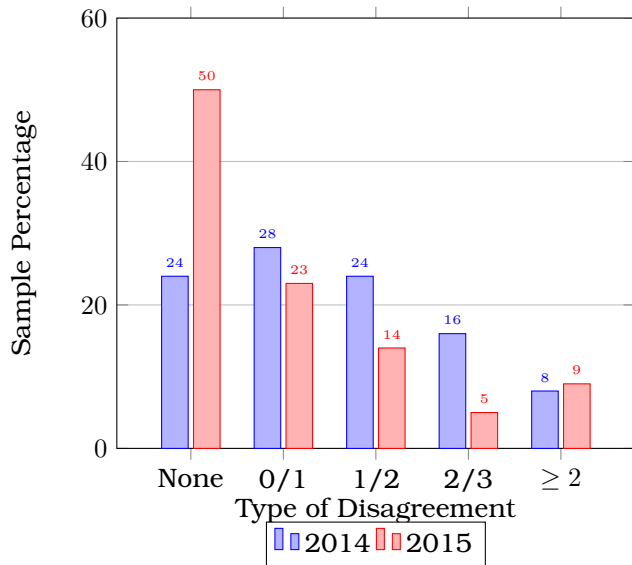
The individual agreement data for the rubric items below highlights an unusually high percentage (25%) of disagreement of two or more for rubric items 2a and 3b, which clearly contributes to the low value of the AC2 statistic for these items (0.36 and 0.38, respectively).



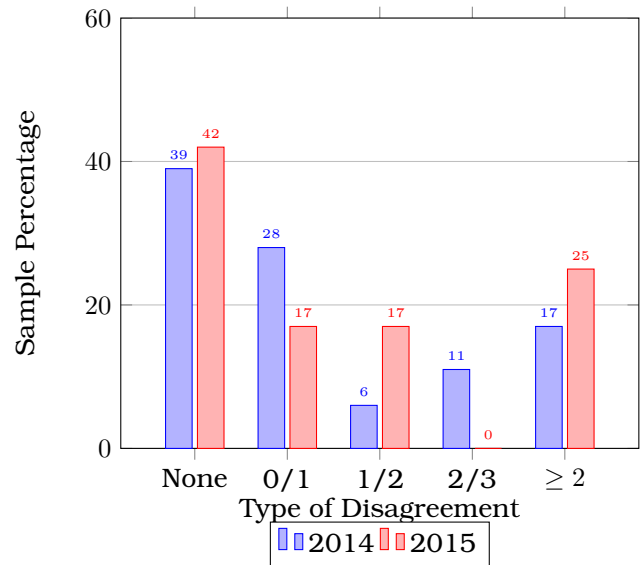
(a) 1: Ability to Access Information, a: Define scope of research question thesis or information needed.



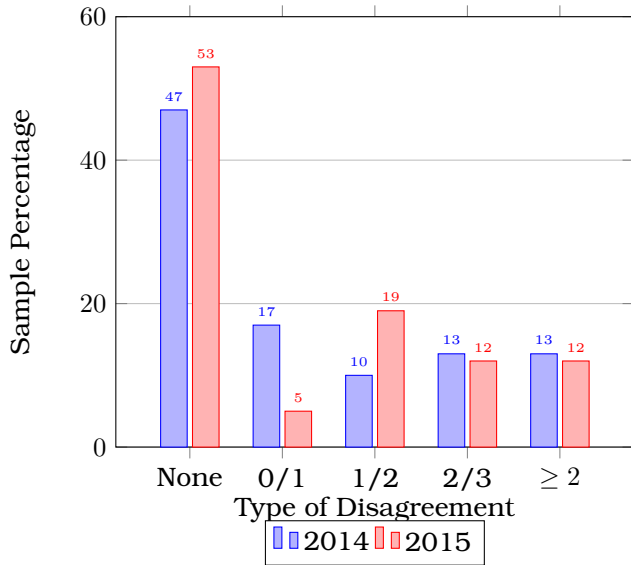
(b) 1: Ability to Access Information, b: Effectively use tools appropriate for a specific task.



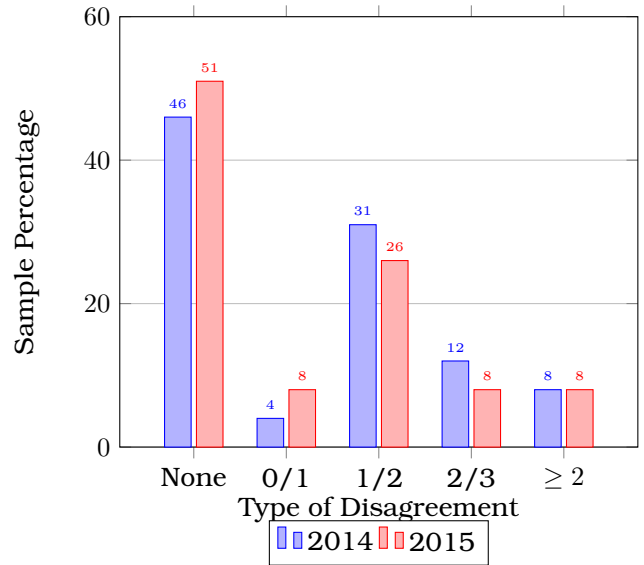
(c) 1: Ability to Access Information, c: Identify key concepts for effective search strategy.



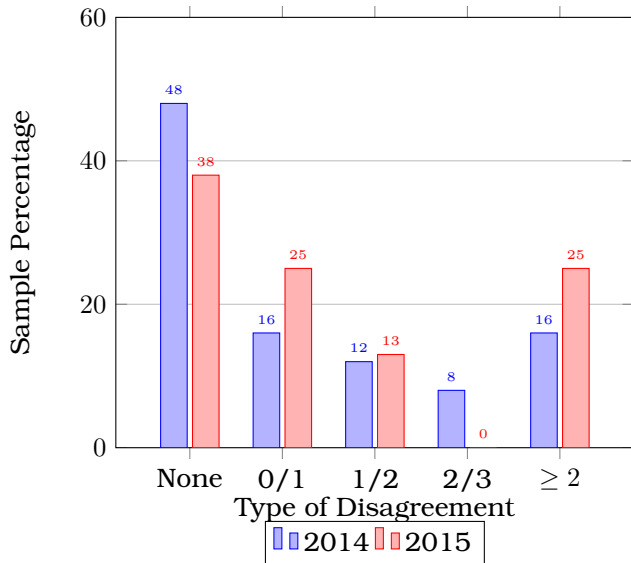
(d) 2: Demonstrate understanding of sources used and information found, a: Differentiates among tertiary, secondary, and primary materials.



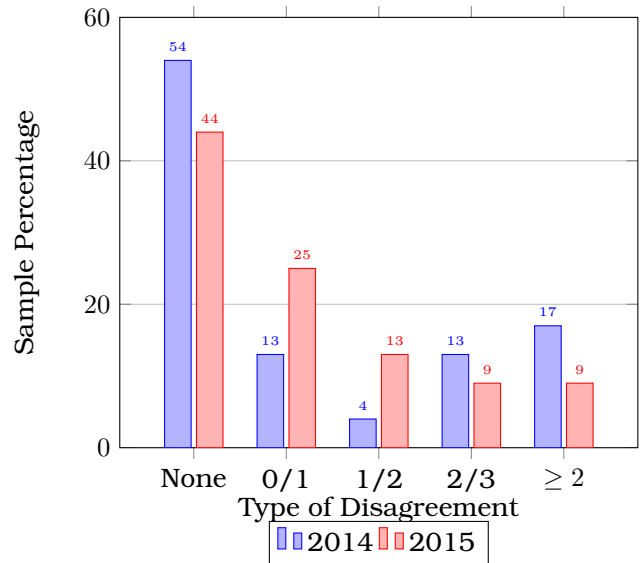
(e) 2: Demonstrate understanding of sources used and information found, b: Selects material or sources, appropriate to the task, considering appropriateness of popular/general sources vs scholarly/academic sources.



(f) 3: Evaluate all information critically, including its sources and authority, a: Use relevant sources to address the research question, topic, or task.

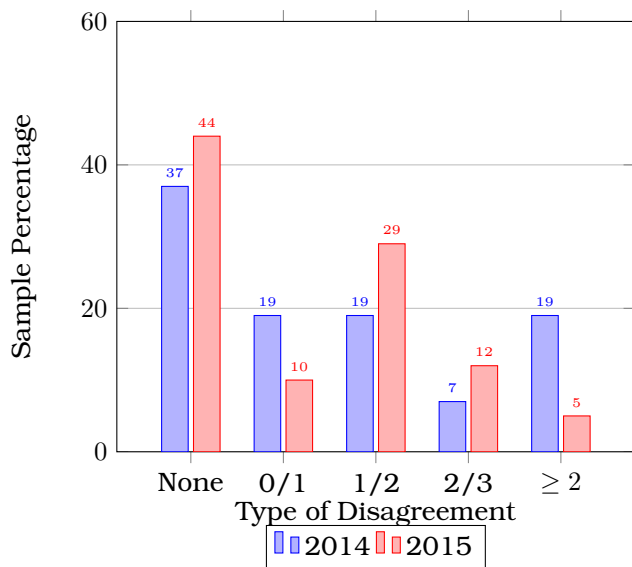


(g) 3: Evaluate all information critically, including its sources and authority, b: Recognize point of view in or quality of material.

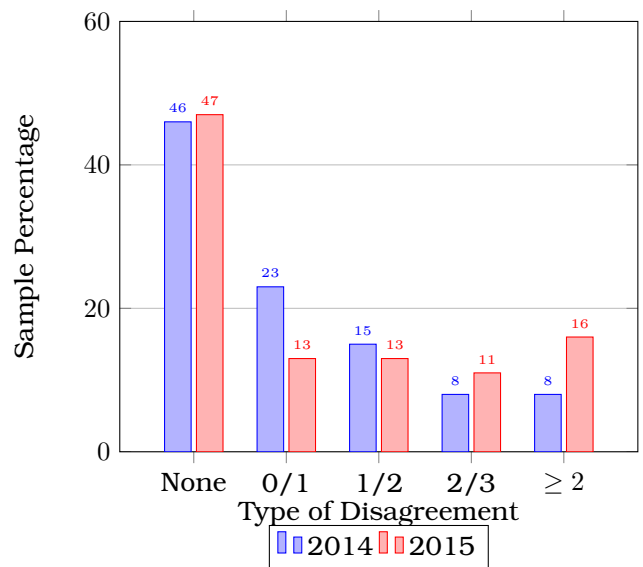


(h) 3: Evaluate all information critically, including its sources and authority, c: Respond to point of view in or quality of material.





(i) 4: Use information responsibly, a: Identify all sources using a consistent citation system.



(j) 4: Use information responsibly, b: Distinguish between common knowledge and ideas requiring attribution.

Figure 34: Type of Disagreement