Investigating Factors Influencing Undergraduates' Performance on General Education

Xiangman Zhao
xiangmaz@andrew.cmu.edu
December 2021

Abstract

Four related questions about finding if various factors, including rater, semester, sex, repeated and rubric, will affect Carnegie Mellon University undergraduates' performances on statistics papers are examined in this paper. The data used is gathered by Dietrich College and includes 91 project papers — referred to as "artifacts"—were randomly sampled from a Fall and Spring section of Freshman Statistics. Different visual plots are used to investigate the distribution of ratings for each rubric. The interclass correlation is used to measure the agreement among raters on each rubric, and linear mixed-effects models are used to investigate what factors are affecting the ratings. Exploratory data analysis shows that the distributions of ratings are different between rubrics and raters, and the final model finds that rubric, rater, semester, and the interaction of rubric and rater are significant factors that influence students' ratings. By understanding how these factors influence students' ratings, Dietrich College can better evaluate students' performances on General Education and how successful the new General Education program is.

1. Introduction

Dietrich College at Carnegie Mellon University is trying to implement a new "General Education" program for all undergraduates, which enables undergraduates to learn new ideas and skills from various disciplines. This program specifies a set of courses and experiences that all undergraduates must take. In order to determine whether the new program is successful, the college hopes to rate student work performed in each of the "General Education" courses each year.

Three raters from across the college are asked to assess the rating work in Freshman Statistics. The paper aims to find the relationship between ratings and different raters and rubrics. Moreover, the relationship between the ratings and other possible factors such as semester, repeated and sex is also investigated.

Four questions related to the research topic are presented below:

1. Is the distribution of ratings for each rubric pretty much indistinguishable from the other rubrics, or are there rubrics that tend to get especially high or low ratings? Is the distribution of ratings given by each rater pretty much indistinguishable from the other raters, or are there raters that tend to give especially high or low ratings?
2. For each rubric, do the raters generally agree on their scores? If not, is there one rater who disagrees with the others? Or do they all disagree?

3. More generally, how are the various factors in this experiment (Rater, Semester, Sex, Repeated, Rubric) related to the ratings? Do the factors interact in any interesting ways?
4. Is there anything else interesting to say about the data?

2. Data

The data for this study provides 91 project papers — referred to as "artifacts"—were randomly sampled from a Fall and Spring section of Freshman Statistics. Three raters from three different departments were asked to rate these artifacts on seven rubrics, as shown in Table 1. The rating scale for all rubrics is shown in Table 2. There are 117 observations and 3 NAs in the dataset. Thirteen artifacts are viewed by all three raters, while the remaining 78 artifacts are viewed by only one rater.

Rubric short name	Rubric full name	Description
RsrchQ	Research Question	Given a scenario, the student generates, critiques, or evaluates a relevant empirical research question.
CritDes	Critique Design	Given an empirical research question, the student critiques or evaluates to what extent a study design convincingly answer that question.
InitEDA	Initial EDA	Given a data set, the student appropriately describes the data and provides initial Exploratory Data Analysis.
SelMeth	Select Method(s)	Given a data set and a research question, the student selects appropriate method(s) to analyze the data.
InterpRes	Interpret Results	The student appropriately interprets the results of the selected method(s).
VisOrg	Visual Organization	The student communicates in an organized, coherent, and effective fashion with visual elements (charts, graphs, tables, etc.).
TxtOrg	Text Organization	The student communicates in an organized, coherent, and effective fashion with text elements (words, sentences, paragraphs, section, and subsection titles, etc.).

Table 1: Rubrics for rating Freshman Statistics projects

Rating	Meaning
1	Student does not generate any relevant evidence.
2	Student generates evidence with significant flaws.
3	Student generates competent evidence; no flaws, or only minor ones.
4	Student generates outstanding evidence; comprehensive and sophisticated.

Table 2: Rating scale used for all rubrics

Table 3 below shows the definitions for all the variables available for the study.

Variable Name	Values	Description
(X)	1, 2, 3, …	Row number in the data set
Rater	1, 2 or 3	Which of the three raters gave a rating
(Sample)	1, 2, 3, …	Sample number
(Overlap)	1, 2, …, 13	Unique identifier for artifact seen by all 3 raters
Semester	Fall or Spring	Which semester the artifact came from
Sex	M or F	Sex or gender of student who created the artifact
RsrchQ	1, 2, 3 or 4	Rating on Research Question
CritDes	1, 2, 3 or 4	Rating on Critique Design
InitEDA	1, 2, 3 or 4	Rating on Initial EDA
SelMeth	1, 2, 3 or 4	Rating on Select Method(s)
InterpRes	1, 2, 3 or 4	Rating on Interpret Results
VisOrg	1, 2, 3 or 4	Rating on Visual Organization
TxtOrg	1, 2, 3 or 4	Rating on Text Organization
Artificial	(text table)	Unique identifier for each artifact
Repeated	0 or 1	1 = this is one of the 13 artifacts seen by all 3 raters

Table 3: Variables definitions

In the study, two dataset files, tall.csv and rating.csv are downloaded for further analysis. These two files contain the same content in different formats. Table 3 shows all the variables in rating.csv, while all seven rubrics are grouped into one variable called rubric in tall.csv.

There are 3 missing values in the dataset, one in Sex, one in CritDes and one in VisOrg. When building the model using the full dataset, missing values in CritDes and VisOrg are dropped. The missing value in Sex is kept in order to include Sex in modeling. When building models on the dataset containing only 13 artifacts viewed by all raters, missing value is not an issue to consider. At last, Table 4 shows the summary statistics of the ratings for each rubric as shown in rating.csv.

Rubric name	Count	Mean	Standard deviation	median	min	max
RsrchQ	114	2.35	0.59	2	1	4
CritDes	114	1.85	0.83	2	1	4
InitEDA	114	2.44	0.70	2	1	4

SelMeth	114	2.05	0.48	2	1	4
InterpRes	114	2.48	0.61	2	1	4
VisOrg	114	2.41	0.68	2	1	4
TxtOrg	114	2.60	0.70	3	1	4

Table 4: Summary statistics of ratings for each rubric

3. Methods

3.1 The distribution of ratings for each rubric and rater

For the first research question, both the full dataset and the dataset only including 13 artifacts viewed by all raters are used to solve the question. Bar plot and counts of the ratings for each rubric are used to compare the distribution of ratings on both datasets. The same process is repeated to compare the distribution of ratings between raters. See Technical Appendix, Pages 3 to 7.

3.2 If raters agree on their ratings for each rubric

For the second research question, a random-intercept model including random intercept with artifact as the grouping variable is fitted for each of the seven rubrics using the dataset only including 13 artifacts viewed by all raters. Rating is the response variable in these models. Interclass correlation for each model is calculated to investigate the correlation between raters on the same artifact. On the dataset including 13 artifacts viewed by all raters, the exact agreement rate is also calculated by making a two-way table of counts of the ratings between each pair of raters for each rubric. The percentages on the main diagonal of the table represent the percentage of times two raters give the same rating for each rubric. On the full dataset, the interclass correlations are calculated using the same method and compared with the interclass correlations on the dataset containing 13 artifacts to see if they agree with each other. See Technical Appendix, Pages 7 to 23.

3.3 Finding significant factors related to the ratings

For the third research question, mixed effects regression is used to find significant factors. There are mainly four steps to solve the question:

1. Using the dataset containing 13 artifacts viewed by all raters, fixed effects are added to seven rubric-specific models trained in research question 2. The optimal model is found using backward elimination and compared with the intercept-only model using a likelihood ratio test to find significant fixed effects for each rubric.
2. Using the full dataset containing all 91 artifacts, fixed effects are added to seven rubric-specific models trained in research question 2. The optimal model is found using backward elimination and compared with the intercept-only model using a likelihood ratio test to find significant fixed effects for each rubric.

3. Continuing with the seven models fitted on the full dataset, interactions between significant fixed effects are added, and random effects are added from all significant fixed effects.
4. In order to explore interactions with Rubric, all significant fixed effects, interactions and random effects from previous steps are added to the "combined model" using Rubric as a random effect grouped by Artifact, using the full dataset. Started with the intercept-only model, fixed effects are chosen using backward elimination, and interactions and random effects are added based on fixed effects chosen. AIC, BIC and likelihood ratio test are used to evaluate models. See Technical Appendix, Pages 24 to 70.

3.4 Investigating additional interesting features of the dataset

For the fourth research question, additional EDAs are tried to explore more on the data and understand the difference between models fitted on the dataset containing only 13 artifacts and the full dataset containing 91 artifacts. Some EDA plots and tables are created to investigate how semester influences ratings. See Technical Appendix, Pages 71 to 85.

4. Results

4.1 The distribution of ratings for each rubric and rater

To answer the first research question, the distributions of ratings for each rubric on the full dataset and on the dataset only containing 13 artifacts are shown in Figure 1 and Figure 2. Counts of four ratings for all artifacts and 13 artifacts are summarized in Table 5 and Table 6. The distribution of ratings for each rater on the full dataset is shown in Figure 3, and the distribution on the dataset containing 13 artifacts and count tables can be found in Part 1 of the Technical Appendix, Pages 3 to 7. It is found that:

- Focusing on the full dataset, rubrics RsrchQ, InitEDA, InterpRes, VisOrg and TxtOrg have similar normal distributions of ratings. Most of ratings are 2 and 3, and they all have very skinny tails, which means very few 1 and 4. Rubrics SelMeth and CritDes have a distribution that is more skewed to the right, which shows that SelMeth and CritDes receive lower ratings compared to other rubrics. CritDes has many ratings of value 1, while SelMeth has many ratings of value 2.
- Focusing on the dataset only containing 13 artifacts, rubrics RsrchQ, InitEDA, SelMeth, InterpRes and VisOrg have a nearly normal distribution of ratings. TxtOrg has a distribution skewing to the left, while CritDes has a distribution skewing to the right.
- Focusing on both the full dataset and each rubric separately, the distribution of ratings for each rater shows that each rater's ratings vary a lot between each rubric. Figure 4 compares the distribution for all artifacts between each rater, it is found that rater 1 and rater 2 tend to give very similar ratings. Most of the ratings given by them are around 2 and 3. Rater 3 tends to give more ratings of 1 and 2, which implies that rater 3 tends to give lower ratings.

More information can be found in Part 1 section (page 3 to 7) of the Technical Appendix.

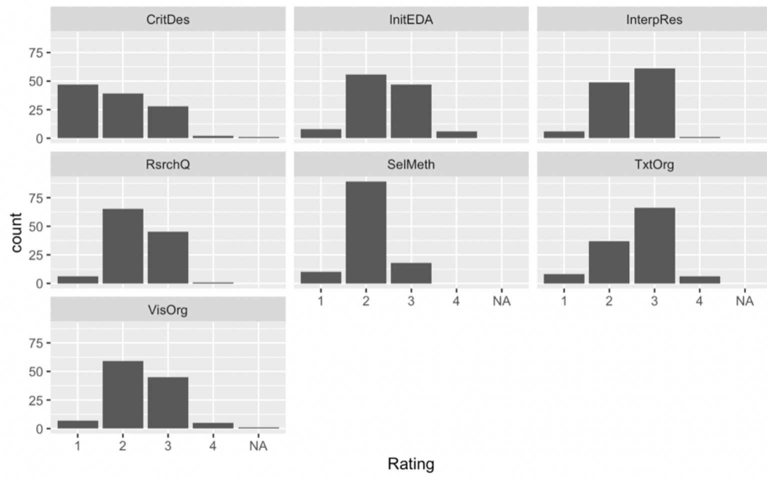


Figure 1: Bar plots of ratings for each rubric for the full dataset

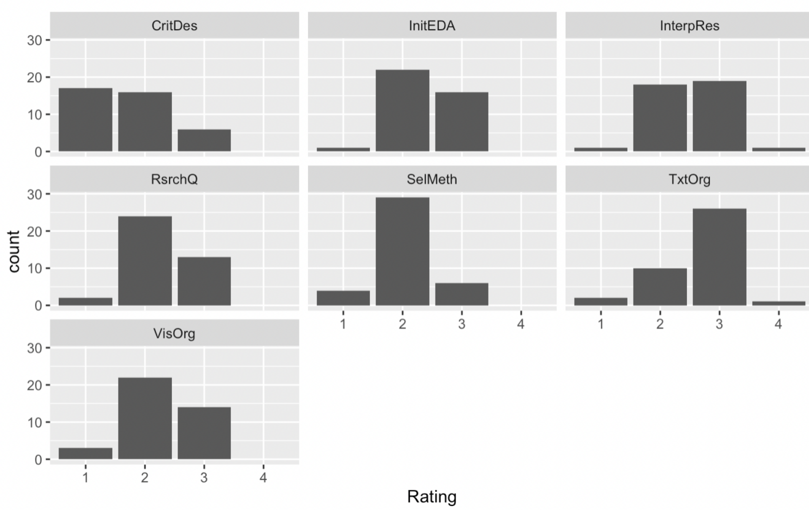


Figure 2: Bar plots of ratings for each rubric for the dataset containing only 13 artifacts

	CritDes	InitEDA	InterpRes	RsrchQ	SelMeth	TxtOrg	VisOrg
Rating 1	47	8	6	6	10	8	7
Rating 2	39	56	49	65	89	37	59
Rating 3	28	47	61	45	18	66	45
Rating 4	2	6	1	1	0	6	5
NA	1	0	0	0	0	0	1

Table 5: Counts of ratings for all artifacts

	CritDes	InitEDA	InterpRes	RsrchQ	SelMeth	TxtOrg	VisOrg
Rating 1	17	1	1	2	4	2	3
Rating 2	16	22	18	24	29	10	22
Rating 3	6	16	19	13	6	26	14
Rating 4	0	0	1	0	0	1	0

Table 6: Counts of ratings for 13 artifacts viewed by all raters

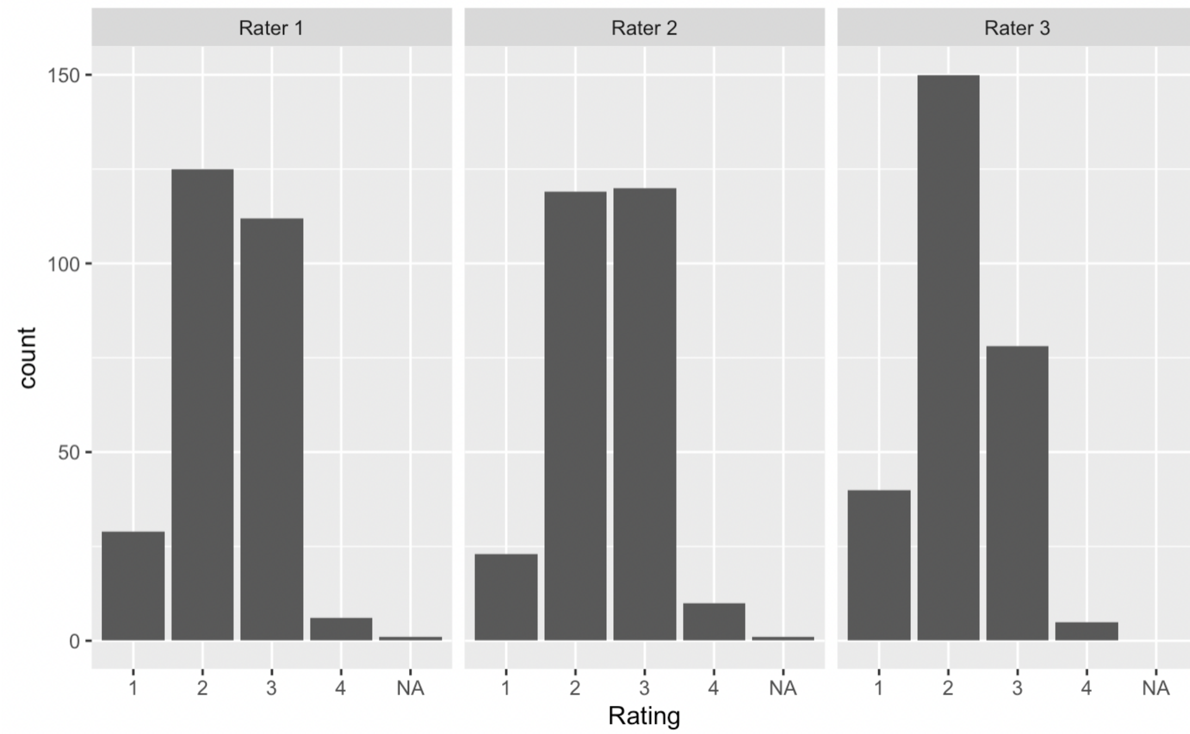


Figure 3: Bar plots of ratings for each rater for the full dataset

4.2 If raters agree on their ratings for each rubric

For the second research question on whether raters agree on their scores, some findings from comparisons between each model's ICC and exact agreement rates between raters in Table 7 are shown below:

1. From interclass correlation, it is found that the correlation between raters on the same artifact for rubrics InitEDA, CritDes, SelMeth and VisOrg is high, which implies that raters' ratings vary a lot for different artifacts for these rubrics. The ICCs for both the full data and common data are low for Rubric RsrchQ, which implies that raters generally do not agree on RsrchQ. ICCs show that there is not a very high agreement rate between raters.
2. From exact agreement rates between raters, raters tend to have consistent ratings on most rubrics, and there is no rater who constantly disagrees with the others. For RsrchQ, the exact agreement rate between rater 1 and rater 2 is the lowest, which means that they don't agree with each other very often. Interestingly, for SelMeth, the exact agreement rate between rater 1 and rater 2 is the highest. Generally, raters' ratings vary with rubrics.
3. Although rubric TxtOrg has relatively a low ICC on both full dataset and common dataset, the exact agreement rate between each pair of raters is still above 0.5. Although rubric SelMeth has high exact agreement rater, the ICC is not that prominent.

Rubric	Interclass correlation		Exact agreement rate		
	Full data	Common data	Rater 1 & 2	Rater 2 & 3	Rater 1 & 3
RsrchQ	0.21	0.19	0.38	0.54	0.77
InitEDA	0.69	0.49	0.69	0.85	0.54
InterpRes	0.22	0.23	0.62	0.62	0.54
CritDes	0.67	0.57	0.54	0.69	0.62
SelMeth	0.47	0.52	0.92	0.69	0.62
TxtOrg	0.19	0.14	0.69	0.54	0.62
VisOrg	0.66	0.59	0.54	0.77	0.77

Table 7: Summary table of ICCs and exact agreement rates

More details on how ICCs and exact agreement rates are calculated can be found in Part 2 section (page 7 to 23) of the Technical Appendix.

4.3 Finding significant factors related to the ratings

The third research question is trying to find significant factors relating to the ratings. Using the data reduced to 13 common artifacts, ANOVA test shows that adding fixed effects does not improve the models (See Technical Appendix Part 3, Pages 24 to 29). The intercept-only model is the best fitted model for each rubric when using the reduced dataset. Next, when using the full dataset, three missing values are removed in order to use the same dataset for each model fitting. The results of the model fitting using the full dataset are different from that using the reduced dataset. For rubrics InitEDA, RsrchQ, and TxtOrg, adding fixed effects does not improve the models. For rubrics CritDes, InterpRes, and VisOrg, adding rater improves the model. For rubric SelMeth, both rater and semester improve the model.

The t-value and the likelihood ratio test show that rater is a significant fixed effect for rubrics CritDes, InterpRes, and VisOrg. Since there is only one fixed effect, there is no need to add interactions. When checking the random intercept of rater grouped by artifact, it is found that there are more random effects than observations, which implies that random effect is not needed. Therefore, for rubrics CritDes, InterpRes, and VisOrg, only rater is added as a fixed effect in the final model.

The model for rubric SelMeth is more complicated as both rater and semester are added as fixed effects. The ANOVA test shows that the interaction of rater and semester is not significant. Similarly, there are more random effects than observations when adding the random intercept of rater grouped by artifact and the random intercept of semester grouped by artifact, which implies that these two random effects are not needed. Therefore, the final model for rubric SelMeth only includes rater and semester as fixed effects.

In addition to build rubric specific models, a combined model including rubric as a single variable is also created to examine interactions with rubric. The starting model includes rubric as a random effect grouped by artifact. The final combined model includes rater, semester, rubric, and the interaction of rater and rubric as fixed effects, as well as rater and rubric as random effects. For more details, see Technical Appendix Part 3, Pages 52 to 70. Table 8 summarizes the coefficients of fixed effects for the final model for each rubric and the combined model. Coefficients of random effects can be found in Technical Appendix Part 3, Pages 64 to 70.

	Rubric specific Models							Combined Model
Fixed effects	RsrchQ	CritDes	InitEDA	SelMeth	InterpRes	VisOrg	TxtOrg	--
Intercept	2.35	--	2.44	--	--	--	2.59	1.76
Rater1	--	1.69	--	2.25	2.70	2.38	--	--
Rater2	--	2.11	--	2.23	2.59	2.65	--	0.37
Rater3	--	1.89	--	2.03	2.14	2.28	--	0.20
SemesterS19	--	--	--	-0.36	--	--	--	-0.16
RubricInitEDA	--	--	--	--	--	--	--	0.74
RubricInterpRes	--	--	--	--	--	--	--	0.99
RubricRsrchQ	--	--	--	--	--	--	--	0.73
RubricSelMeth	--	--	--	--	--	--	--	0.41
RubricTxtOrg	--	--	--	--	--	--	--	1.02
RubricVisOrg	--	--	--	--	--	--	--	0.65
Rater2: RubricEDA	--	--	--	--	--	--	--	-0.30
Rater3: RubricInitEDA	--	--	--	--	--	--	--	-0.29
Rater2: RubricInterpRes	--	--	--	--	--	--	--	-0.51
Rater3: RubricInterpRes	--	--	--	--	--	--	--	-0.71
Rater2: RubricRsrchQ	--	--	--	--	--	--	--	-0.49
Rater3: RubricRsrchQ	--	--	--	--	--	--	--	-0.32
Rater2: RubricSelMeth	--	--	--	--	--	--	--	-0.39
Rater3: RubricSelMeth	--	--	--	--	--	--	--	-0.39
Rater2: RubricTxtOrg	--	--	--	--	--	--	--	-0.55
Rater3: RubricTxtOrg	--	--	--	--	--	--	--	-0.44
Rater2: RubricVisOrg	--	--	--	--	--	--	--	-0.10
Rater3: RubricVisOrg	--	--	--	--	--	--	--	-0.28

Table 8: Coefficients of Fixed Effects

Interpretations for the intercept-only models for rubrics RsrchQ, InitEDA and TxtOrg:
These models only have intercepts which represent the overall mean of ratings for each rubric. Take RsrchQ as an example, the overall mean of ratings is 2.35, and the mean of ratings for artifact 100 is 2.35 – 0.07 = 2.28, which means that the rating for RsrchQ of artifact 100 is close to the overall mean of ratings for all artifacts.

Interpretations for models including rater as a fixed effect for rubrics CritDes, InterpRes and VisOrg:
Different raters give different ratings, so the interpretation depends on both the rubric and the rater. Take rubric CritDes as an example, the overall mean of ratings given by rater 1 is 1.69, and

the mean of ratings given by rater 1 for artifact 100 is 1.69 + 0.84 = 2.53, which implies that the rating for CritDes of artifact 100 given by rater 1 is higher than the mean of overall ratings given by rater 1 for rubric CritDes of all artifacts.

Interpretations for models including rater and semester as fixed effects for rubrics SelMeth:

The rating depends on rater and semester. The overall mean of ratings given by rater 1 during fall semester 2019 is 2.25, and the mean rating given by rater 1 for artifact 100 during fall semester 2019 is 2.25 + 0.34 = 2.59, which implies that the rating for SelMeth of artifact 100 given by rater 1 during fall semester 2019 is slightly higher than the mean of overall ratings given by rater 1 for rubric SelMeth of all artifacts during fall semester 2019.

Interpretations for the combined model:

The interpretation will be broken down into three parts based on variables.

For rater added both as a fixed effect and as a random effect grouped by artifact, there is a kind of interaction between rater and artifact. Each rater's rating on each artifact differs from the fixed effect alone by a small random effect that depends on the artifact. Take artifact 100 as an example, the mean rating given by rater 1 is 1.76 + 0.04 = 1.80, which implies that the rating given by rater 1 is almost the same as the overall rating given by rater 1.

For rater, rubric and the interaction of rater and rubric added as fixed effects, there is a kind of interaction between rater and rubric. Each rater uses each rubric in a way that is not the same, or even parallel to, other rater's rubric usage. Take RsrchQ as an example, the mean of ratings given by rater 2 for rubric RsrchQ is 0.37 + 0.73 – 0.49 = 0.61 higher, compared to the mean rating given by rater 1 for rubric CritDes.

For rubric added both as a fixed effect and as a random effect grouped by artifact, there is a kind of interaction between rubric and artifact. There are different average scores on each rubric, but the rubric averages also vary a bit from one artifact to another, by a small random effect depending on artifact. Take RsrchQ as an example, the mean rating for RsrchQ of artifact 100 is 0.73 – 0.17 = 0.56, which implies that the rating given for RsrchQ of artifact 100 is lower than the overall rating given for rubric RsrchQ.

All of the random effect coefficients and more details on comparing models can be found in Technical Appendix Part 3, Pages 24 to 70.

4.4 Investigating additional interesting features of the dataset

For the fourth research question, two bar plots (Technical Appendix, Pages 70 -71) shows that the ratings distribution between these two semesters are not that different, but semester is a significant variable for the model. One reason is that the dataset is not evenly split. There are 580 observations in the fall semester, which is more than twice of 237 observations in the spring semester, which may cause the difference between ratings in these two semesters. In order to better explore the difference of ratings between two semester, Table 9 shows the percent of each rating given in fall and spring semester, and Table 10 shows a more detailed percent of each rating by rubric. Overall, raters tend to give higher ratings in the fall semester, and the percent of rating of 1 in the fall semester is especially less than that in the spring semester. Table 10 shows that rubrics InterpRes, CritDes and SelMeth probably contribute more to the overall lower ratings in the spring semester. The rating of 1 given to InterpRes in the spring is 9% higher than that given in the fall semester, while the rating of 1 given to CritDes is more than 15% higher than that given in the fall semester. The ratings for rubric SelMeth are especially low in the spring semester, as there are no ratings of 3 or 4 given, and there are almost 13% more ratings of 1 given. The large difference in the ratings between fall semester and spring semester explains why semester is chosen as a fixed effect in the model for rubric SelMeth. The difference in the overall ratings also explains why semester is a significant factor in the final model.

Table 11 summarizes the ratings given by each rater between two semesters. Rater 1 and rater 2 have nearly consistent ratings, while rater 3 gives much lower ratings in the spring semester compared to the fall semester. Rater 3 gives 24.6% more ratings of 1 in the spring semester, which is probably the main reason why ratings in the spring is much lower than that in the fall.

More information on how these percentages are calculated can be found in Technical Appendix Part 4, Pages 70 to 85.

	Rating 1	Rating 2	Rating 3	Rating 4
Fall	8.6%	50.3%	38.8%	2.3%
Spring	18.2%	44.1%	34.3%	3.4%

Table 9: Percentage table of rating in two semesters

	Fall				Spring			
Rubric	Rating 1	Rating 2	Rating 3	Rating 4	Rating 1	Rating 2	Rating 3	Rating 4
RsrchQ	8.8%	41.2%	50%	0%	8.8%	41.2%	50%	0%
InitEDA	6.1%	47.6%	42.7%	3.6%	8.8%	50%	32.4%	8.8%
InterpRes	2.5%	46.3%	50%	1.2%	11.8%	32.3%	55.9%	0%
CritDes	36.6%	37.8%	24.4%	1.2%	53.1%	25.0%	21.8%	3.1%
SelMeth	4.9%	74.4%	20.7%	0%	17.6%	82.4%	0%	0%
TxtOrg	4.9%	30.4%	59.8%	4.9%	11.8%	35.3%	47.1%	5.8%
VisOrg	1.3%	53.7%	41.2%	3.8%	17.6%	44.1%	32.4%	5.9%

Table 10: Percentage table of rating by rubrics in two semesters

	Fall				Spring			
Rubric	Rating 1	Rating 2	Rating 3	Rating 4	Rating 1	Rating 2	Rating 3	Rating 4
Rater 1	9.6%	46.0%	42.2%	2.2%	13.1%	45.2%	39.3%	2.4%
Rater 2	8.1%	43.9%	44.9%	3.1%	9.3%	42.7%	42.7%	5.3%
Rater 3	7.9%	61.4%	29.1%	1.6%	32.5%	44.2%	20.8%	2.5%

Table 11: Percentage table of rating by raters in two semesters

5. Discussions

When Dietrich College tries to evaluate how successful the new general education program is, they should consider the effect of rater, rubric, semester, and artifact on students' ratings. The interaction of rater and rubric also influences students' ratings. The result that ratings depend on artifact is expected, the interaction of rater and rubric, as well as the interaction of rater and artifact, are problematic, which shows that there is not a consistent rating between rubrics and raters. More details are explained in the sub-sections below.

5.1 The distribution of ratings for each rubric and rater

The distributions of ratings between different rubrics and raters are not largely different. Comparing between the full dataset and the dataset containing 13 artifacts viewed by all raters, the distribution of ratings for each rubric is similar, which shows that variable repeated is not very significant. Some rubrics such as CritDes and SelMeth tend to have lower ratings than the others, which might be due to the higher level of difficulty or the poor quality of teaching in these two rubrics. Focusing on the distribution of ratings for each rater, each rater's rating varies a lot between rubrics, but rater 3 tends to give lower ratings. Some possible reasons are that rater 3 grades more strictly than the others, and rater 3 might understand the rubric differently from the other two raters, which may require more future trainings for more consistent ratings.

5.2 If raters agree on their ratings for each rubric

For the second research question, it is found that all raters agree with each other for most of rubrics. Especially for rubrics InitEDA, CritDes, SelMeth and VisOrg, the correlation between raters' ratings on the same artifact is high, which means that raters are consistent in their ratings for these rubrics. Raters might need more trainings for consistent ratings on rubrics RsrchQ, InterpRes and TxtOrg. From exact agreement rates between raters, raters' ratings vary by rubrics, which is expected as the level of difficulty of each rubric is different. The result for the second question corresponds with the result in the first research question that ratings largely depend on the rubric instead of raters.

5.3 Finding significant factors related to the ratings

When focusing on the small dataset containing only 13 artifacts viewed by all raters, intercept-only models with not fixed effect are chosen. The result shows that the ratings are not affected by raters, sex, or other variables. When using the full dataset, some rubrics have raters and semester as fixed effects, while rubrics RsrchQ, InitEDA and TxtOrg still choose intercept-only models. For other rubrics, rater and semester will affect the rating, which indicates problematic gradings and needs further improvements in the training. For the final model on the full dataset, rater, semester, rubric and the interaction of rater and rubric are chosen as the significant fixed effects, and rater and rubric are also chosen as the random effects. It is expected that rubric is chosen as a fixed effect as some rubrics are straightforward and tend to have higher ratings like InterpRes and TxtOrg, and rubrics like SelMeth is probably harder and receives generally lower ratings. It is problematic that semester and rater are fixed effects, which shows that the ratings are not consistent between different raters and semesters. It is also problematic that the interaction of rater and rubric is significant, and rater is chosen as a random effect grouped by artifact, which implies that raters don't interpret each rubric and each artifact in the same way. One possible explanation is that raters come from different departments and probably focus on different perspectives in their ratings on each artifact, which again requires more trainings for more consistent ratings for fairness. More exploration on how semester affects the rating is discussed in the fourth research question.

5.4 Investigating additional interesting features of the dataset

For the last question, it is shown that ratings given in the fall are higher than ratings given in the spring. The result is kind of problematic as freshmen tend to receive lower grades in the first semester as most of them are still transferring from high school teaching modes to college teaching modes. In the spring semester, it is assumed that freshmen are used to the college life and should perform better. There are more ratings of 1 in the spring, and rubrics SelMeth and CritDes receive especially more lower ratings in the spring. It is also found that rater 3 tends to give nearly 25% more ratings of 1 in the spring semester, which contributes to the lower ratings in the spring. The result shows that rater 3 has a very inconsistent pattern of ratings between spring and fall semester. Dietrich college should talk with rater 3 to understand how he/she grades and have more training sessions for rater 3 to help him/her grade more consistently.

5.5 Limitations and future improvements

Some strengths of the modeling part are that both full dataset and the dataset containing only 13 artifacts viewed by all raters are taken into considerations, which shows mores features of the dataset. In addition, results from the EDA help to explain why some fixed effects are chosen in the final model, which helps the audience to understand and interpret the final model. There are still some limitations with the final model. First, the dataset is small and contains some missing values including the missing input values for sex. When fitting the models, these missing values are just removed. It might be better to consider a more comprehensive way to deal with these missing values. Second, the dataset is kind of small, especially for the reduced dataset of artifacts viewed by all raters. The number of observations in the fall semester is also much larger than that in the spring semester, which might cause some biases in the result. For future improvements, some other models can be fitted to see if the significant variables will be any different. For example, how about treating the response variable as a categorical variable and fitting a classification model to if the significant variables will be different. Also, more artifacts should be collected to build a larger dataset. More variables that may affect students' performances such as major can also be added for future analysis. These future steps can be taken to understand the dataset better.

All in all, rubric, rater, semester, and the interaction of rubric and rater are found to be the fixed effects in the final model. These factors are problematic, which implies unfairness in the rating system and suggests more trainings for raters. The suggestion for Dietrich college is to always consider these factors when evaluating students' ratings to make better decisions to help students to achieve successes at school.

6. References

Sheather, S.J. (2009), *A Modern Approach to Regression with R.* New York: Springer Science + Business Media LLC.

Technical Appendix

Technical Appendix

This is an R Markdown Notebook. When you execute code within the notebook, the results appear beneath the code.

Try executing this chunk by clicking the *Run* button within the chunk or by placing your cursor inside it and pressing *Cmd+Shift+Enter*.

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.5     v dplyr   1.0.7
## v tidyr   1.1.4     v stringr 1.4.0
## v readr   2.0.1     v forcats 0.5.1
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(kableExtra)
```

```
##
## Attaching package: 'kableExtra'
```

```
## The following object is masked from 'package:dplyr':
##
##     group_rows
```

```
library(GGally)
```

```
## Registered S3 method overwritten by 'GGally':
##   method from
##   +.gg   ggplot2
```

```
library(grid)
library(gridExtra)
```

```
##
## Attaching package: 'gridExtra'
```

```
## The following object is masked from 'package:dplyr':
##
##     combine
```

```
library(ggplotify)
library(reshape2)
```

```
##
## Attaching package: 'reshape2'
```

```
## The following object is masked from 'package:tidyr':
##
##     smiths
```

```
library(plyr)
```

```
## ------------------------------------------------------------------------------

## You have loaded plyr after dplyr - this is likely to cause problems.
## If you need functions from both plyr and dplyr, please load plyr first, then dplyr:
## library(plyr); library(dplyr)

## ------------------------------------------------------------------------------

##
## Attaching package: 'plyr'

## The following objects are masked from 'package:dplyr':
##
##     arrange, count, desc, failwith, id, mutate, rename, summarise,
##     summarize

## The following object is masked from 'package:purrr':
##
##     compact
```

```
library(lme4)
```

```
## Loading required package: Matrix

##
## Attaching package: 'Matrix'

## The following objects are masked from 'package:tidyr':
##
##     expand, pack, unpack
```

```
library(arm)
```

```
## Loading required package: MASS

##
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':
##
##     select

##
## arm (Version 1.12-2, built: 2021-10-15)

## Working directory is /Users/zhaoxiangman/Desktop/36-617
```

```
library(performance)
```

```
##
## Attaching package: 'performance'

## The following object is masked from 'package:arm':
##
##     display
```

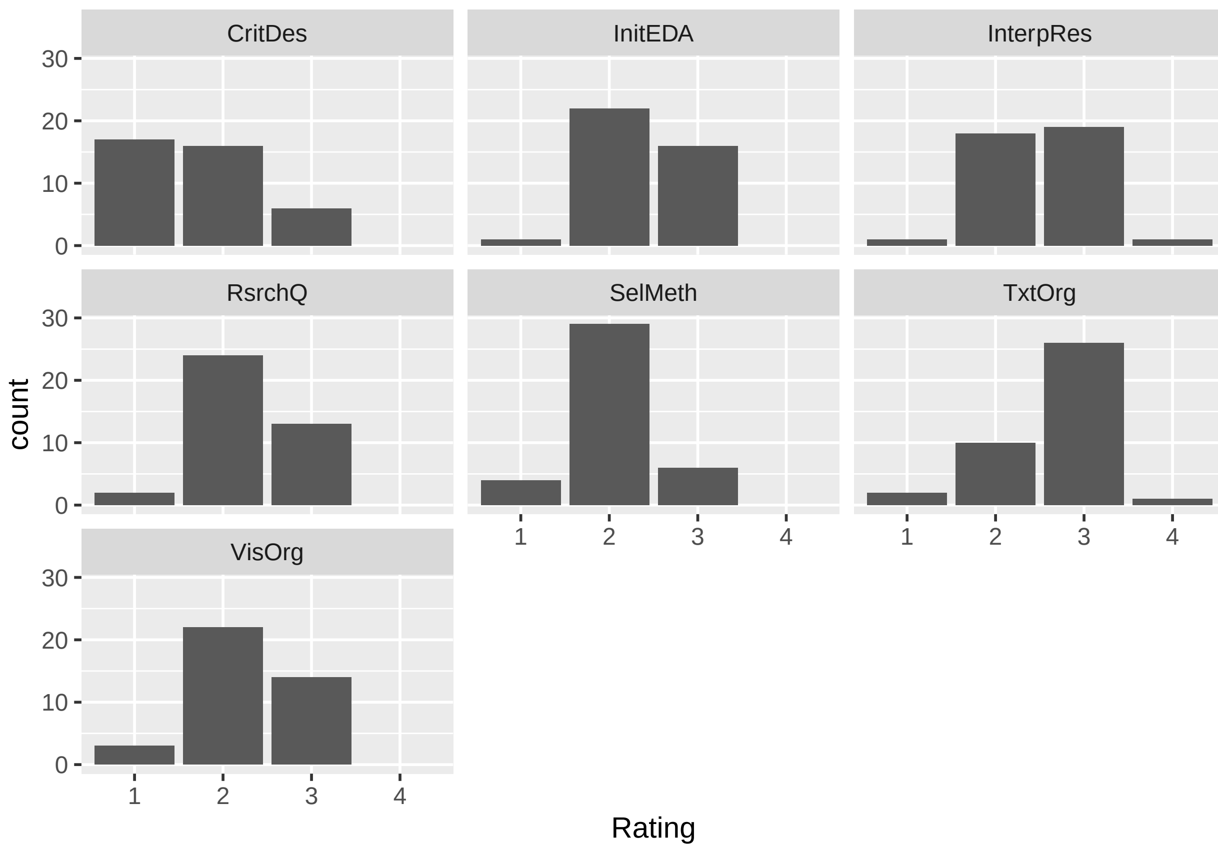
```
cdi <- read.table("/Users/zhaoxiangman/Desktop/36-617/cdi.dat", header = TRUE)
```

Part 1

```r
ratings <- read.csv("/Users/zhaoxiangman/Desktop/36-617/ratings.csv")
tall <- read.csv("/Users/zhaoxiangman/Desktop/36-617/tall.csv")
```

```r
tall$Rating <- factor(tall$Rating,levels=1:4)
for (i in unique(tall$Rubric)) {
  ratings[,i] <- factor(ratings[,i],levels=1:4)
}
tall$Sex[nchar(tall$Sex)==0] <- "--"
ratings.13 <- ratings[grep("O",ratings$Artifact),]
tall.13 <-  tall[grep("O",tall$Artifact),]
```

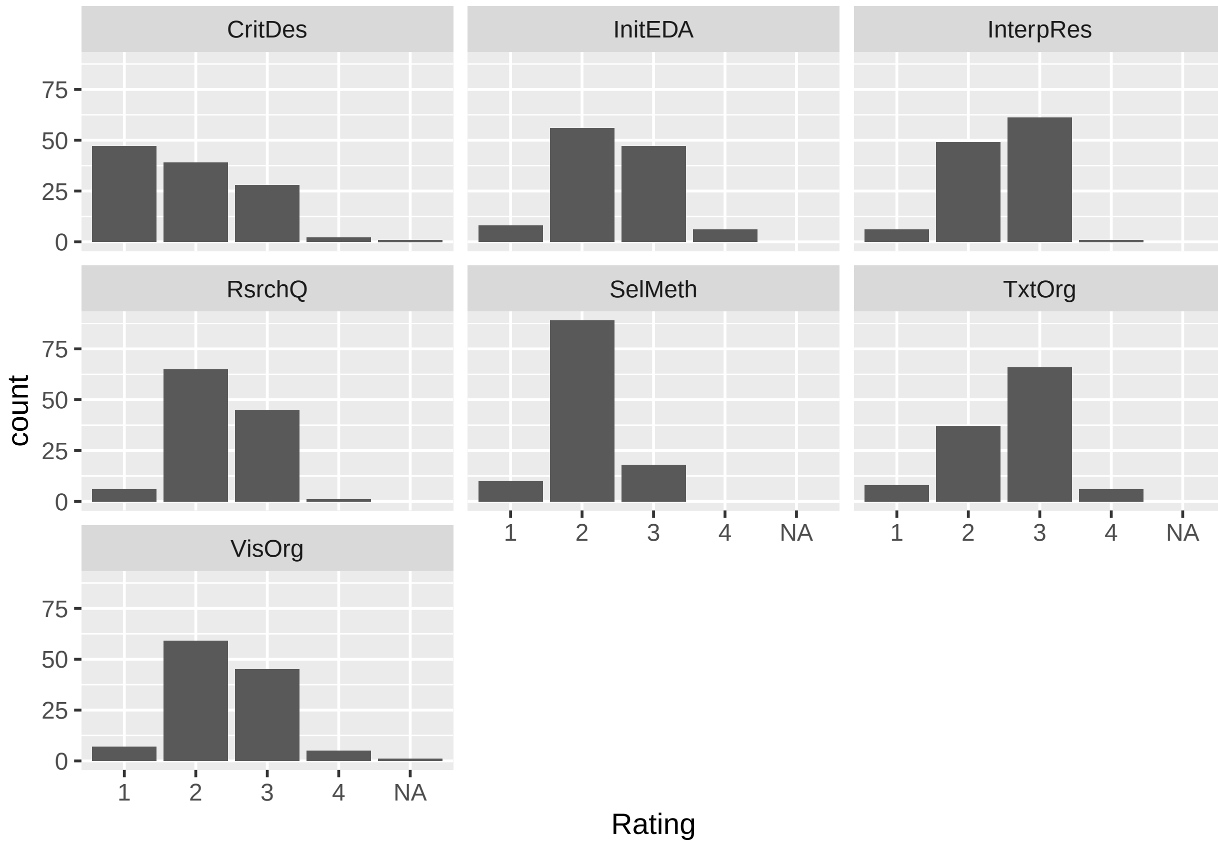
```r
g <- ggplot(tall.13,aes(x = Rating)) +
  facet_wrap( ~ Rubric) +
  geom_bar()
g
```




```r
tmp <- data.frame(lapply(split(tall.13$Rating,tall.13$Rubric),summary))
row.names(tmp) <- paste("Rating",1:4)

tmp
```

```
##          CritDes InitEDA InterpRes RsrchQ SelMeth TxtOrg VisOrg
## Rating 1      17       1         1      2       4      2      3
## Rating 2      16      22        18     24      29     10     22
## Rating 3       6      16        19     13       6     26     14
## Rating 4       0       0         1      0       0      1      0
```

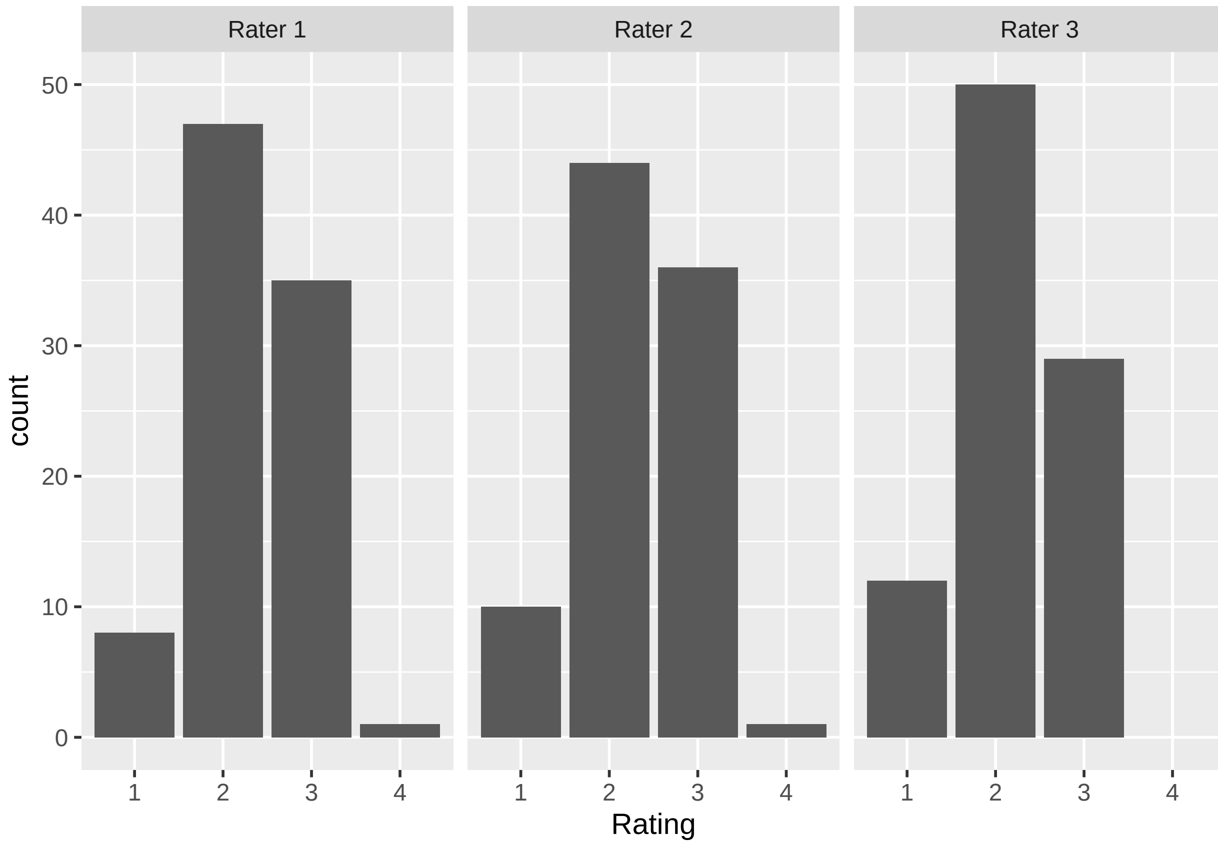
```r
#
## Barplots for full data set
g <- ggplot(tall,aes(x = Rating)) +
  facet_wrap( ~ Rubric) +
  geom_bar()

g
```




```r
##
## Table of counts again.  A bit pesky since there are NA's...
tmp0 <- lapply(split(tall$Rating,tall$Rubric),summary)
tmp <- data.frame(matrix(0,nrow=5,ncol=7))  ## seven rubrics...
names(tmp) <- names(tmp0)
row.names(tmp) <- c(paste("Rating",1:4),"<NA>")
for (i in names(tmp0)) {
  tmp[,i] <- tmp[,i] + c(tmp0[[i]],0)[1:5]
}

tmp
```

```
##          CritDes InitEDA InterpRes RsrchQ SelMeth TxtOrg VisOrg
## Rating 1      47       8         6      6      10      8      7
## Rating 2      39      56        49     65      89     37     59
## Rating 3      28      47        61     45      18     66     45
## Rating 4       2       6         1      1       0      6      5
## <NA>           1       0         0      0       0      0      1
```

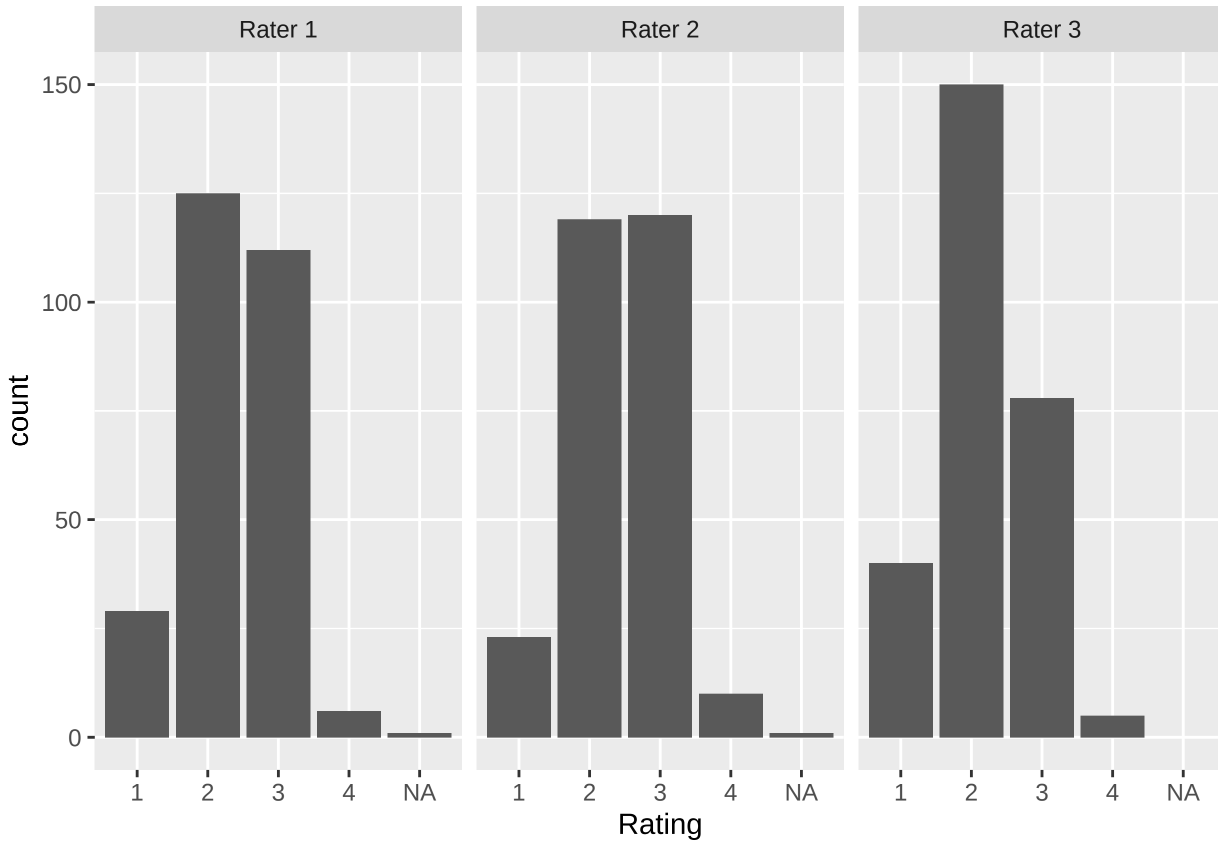
```r
##
## Needed to make the title of each facet more human-readable...
rater.name <- function(x) { paste("Rater",x) }

##
## Barplots for reduced data...
g <- ggplot(tall.13,aes(x = Rating)) +
  facet_wrap( ~ Rater, labeller=labeller(Rater=rater.name)) +
  geom_bar()

g
```




```r
##
## Corresponding table of counts...
tmp <- data.frame(lapply(split(tall.13$Rating,tall.13$Rater),summary))
row.names(tmp) <- paste("Rating",1:4)
names(tmp) <- paste("Rater",1:3)

tmp
```

```
##          Rater 1 Rater 2 Rater 3
## Rating 1       8      10      12
## Rating 2      47      44      50
## Rating 3      35      36      29
## Rating 4       1       1       0
```

```
##
## Barplots for full data...
g <- ggplot(tall,aes(x = Rating)) +
  facet_wrap( ~ Rater, labeller=labeller(Rater=rater.name)) +
  geom_bar()

g
```




```
##
## Corresponding table of counts...
tmp0 <- lapply(split(tall$Rating,tall$Rater),summary)
tmp <- data.frame(matrix(0,nrow=5,ncol=3))  ## three raters...
names(tmp) <- names(tmp0)
row.names(tmp) <- c(paste("Rating",1:4),"<NA>")
for (i in names(tmp0)) {
  tmp[,i] <- tmp[,i] + c(tmp0[[i]],0)[1:5]
}
names(tmp) <- paste("Rater",1:3)
tmp
```

```
##          Rater 1 Rater 2 Rater 3
## Rating 1      29      23      40
## Rating 2     125     119     150
## Rating 3     112     120      78
## Rating 4       6      10       5
## <NA>           1       1       0
```

Histograms shows that each Rurbic tends to have different distribution. RsrchQ, InitEDA, InterpRes, VisOrg and TxtOrg have rather similar distributions. Most of the ratings are around 2 for SelMeth, The distribution of CritDes is skewed to the right as there are much more lower ratings for CritDes.

For rater 1, the distribution of ratings on different rubrics are different. Rubrics Research question, InitEDA, VisOrg and TxtOrg all have four ratings and most of the ratings have 2 and 3s. Rater 1's ratings on CritDes question is mostly 1s and the second most ratings are 2, and the least one is 3. For rubric SelMeth, there are only two ratings 2 and 3. There are much more 2s than 3s. The rubric InterpRes also only has two ratings 2 and 3. There are more 3s than 2s.

For rater 2, rubric research question has most ratings 2 and the next most ratings is 3 and the least rating is 1. For rubric CritDes, most ratings are 1, 2, and 3, and 4 is the least. For rubric InitEDA, most ratings are 2 and 3, and there are only 1 and 4. For rubric SelMeth, most of the ratings are 2 and there are much less 1s and 3s. For InterpRes, VisOrg and Txt the most rating is 3 and 2, there are only very few 1 and 4.

Now I will take a closer look at the subset of 13 artifacts that are viewed by all raters. The distribution of ratings is very different from that on the entire dataset. Each rater is very different from their ratings. Ratings for rubric rsrchQ, ratings for Rubric InitEDA and ratinsg for rubric Visorg are very similar to each other, as rating 2 has the most counts and 3 has less count, and there are only a few 1s. Rubric CritDes has lower ratings as there are more 1s and 2s. Rubric InterpRes and TxtOrg both have 4 ratings and most are centered aroung 2 and 3.

Part 2

```
tall <- tall %>% na.omit()
summary(tall)
```

```
##        X            Rater         Artifact            Repeated
##  Min.   :  1   Min.   :1.000   Length:817         Min.   :0.0000
##  1st Qu.:206   1st Qu.:1.000   Class :character   1st Qu.:0.0000
##  Median :410   Median :2.000   Mode  :character   Median :0.0000
##  Mean   :410   Mean   :2.001                      Mean   :0.3341
##  3rd Qu.:614   3rd Qu.:3.000                      3rd Qu.:1.0000
##  Max.   :819   Max.   :3.000                      Max.   :1.0000
##    Semester             Sex               Rubric            Rating
##  Length:817         Length:817         Length:817         1: 92
##  Class :character   Class :character   Class :character   2:394
##  Mode  :character   Mode  :character   Mode  :character   3:310
##                                                           4: 21
##
##
```

```
common <- tall[grep("O",tall$Artifact),]
```

```
dim(common)
```

```
## [1] 273   8
```

```
RsrchQ.ratings <- common[common$Rubric=="RsrchQ",]
CritDes.ratings <- common[common$Rubric == "CritDes",]
InitEDA.ratings <- common[common$Rubric == "InitEDA",]
SelMeth.ratings <- common[common$Rubric == "SelMeth",]
InterpRes.ratings <- common[common$Rubric == "InterpRes",]
VisOrg.ratings <- common[common$Rubric == "VisOrg",]
TxtOrg.ratings <- common[common$Rubric == "TxtOrg",]
```

```r
summary(tall)
```

```
##        X             Rater         Artifact            Repeated
##  Min.   :  1   Min.   :1.000   Length:817         Min.   :0.0000
##  1st Qu.:206   1st Qu.:1.000   Class :character   1st Qu.:0.0000
##  Median :410   Median :2.000   Mode  :character   Median :0.0000
##  Mean   :410   Mean   :2.001                      Mean   :0.3341
##  3rd Qu.:614   3rd Qu.:3.000                      3rd Qu.:1.0000
##  Max.   :819   Max.   :3.000                      Max.   :1.0000
##    Semester             Sex                Rubric          Rating
##  Length:817         Length:817         Length:817         1: 92
##  Class :character   Class :character   Class :character   2:394
##  Mode  :character   Mode  :character   Mode  :character   3:310
##                                                           4: 21
##
##
```

```r
mod1 <- lmer(as.numeric(Rating) ~ 1 + (1|Artifact), data=RsrchQ.ratings)
mod2 <- lmer(as.numeric(Rating) ~ 1 + (1|Artifact), data=CritDes.ratings)
mod3 <- lmer(as.numeric(Rating) ~ 1 + (1|Artifact), data=InitEDA.ratings)
mod4 <- lmer(as.numeric(Rating) ~ 1 + (1|Artifact), data=SelMeth.ratings)
mod5 <- lmer(as.numeric(Rating) ~ 1 + (1|Artifact), data=InterpRes.ratings)
mod6 <- lmer(as.numeric(Rating) ~ 1 + (1|Artifact), data=VisOrg.ratings)
mod7 <- lmer(as.numeric(Rating) ~ 1 + (1|Artifact), data=TxtOrg.ratings)
```

```r
RsrchQ.full <- tall[tall$Rubric=="RsrchQ",]
CritDes.full <- tall[tall$Rubric == "CritDes",]
InitEDA.full <- tall[tall$Rubric == "InitEDA",]
SelMeth.full <- tall[tall$Rubric == "SelMeth",]
InterpRes.full <- tall[tall$Rubric == "InterpRes",]
VisOrg.full <- tall[tall$Rubric == "VisOrg",]
TxtOrg.full <- tall[tall$Rubric == "TxtOrg",]
```

```r
mod1_1 <- lmer(as.numeric(Rating) ~ 1 + (1|Artifact), data=RsrchQ.full)
mod2_1 <- lmer(as.numeric(Rating) ~ 1 + (1|Artifact), data=CritDes.full)
mod3_1 <- lmer(as.numeric(Rating) ~ 1 + (1|Artifact), data=InitEDA.full)
mod4_1 <- lmer(as.numeric(Rating) ~ 1 + (1|Artifact), data=SelMeth.full)
mod5_1 <- lmer(as.numeric(Rating) ~ 1 + (1|Artifact), data=InterpRes.full)
mod6_1 <- lmer(as.numeric(Rating) ~ 1 + (1|Artifact), data=VisOrg.full)
mod7_1 <- lmer(as.numeric(Rating) ~ 1 + (1|Artifact), data=TxtOrg.full)
```

```r
icc(mod1)
```

```
## # Intraclass Correlation Coefficient
##
##      Adjusted ICC: 0.189
##   Conditional ICC: 0.189
```

```r
icc(mod2)
```

```
## # Intraclass Correlation Coefficient
##
##      Adjusted ICC: 0.573
##   Conditional ICC: 0.573
```

```r
icc(mod3)
```

```
## # Intraclass Correlation Coefficient
##
##      Adjusted ICC: 0.493
##   Conditional ICC: 0.493
```

```
icc(mod4)
```

```
## # Intraclass Correlation Coefficient
##
##      Adjusted ICC: 0.521
##   Conditional ICC: 0.521
```

```
icc(mod5)
```

```
## # Intraclass Correlation Coefficient
##
##      Adjusted ICC: 0.230
##   Conditional ICC: 0.230
```

```
icc(mod6)
```

```
## # Intraclass Correlation Coefficient
##
##      Adjusted ICC: 0.592
##   Conditional ICC: 0.592
```

```
icc(mod7)
```

```
## # Intraclass Correlation Coefficient
##
##      Adjusted ICC: 0.143
##   Conditional ICC: 0.143
```

> Just looking at the subset of artifact viewed by all raters, and looking at each rubric separately, Rubric CritDes, InitEDA, SelMeth and VisOrg have high ICC which means high correlation between raters and ratings on the same artifact, which means that the raters are consistent with one another in how they rate. Rubric RsrchQ, InterpRes and TxtOrg all have lower ICC which means low correlation between raters and ratings on the same artifact, which means that the raters are not consistent with one another in how they rate.

```
icc(mod1_1)
```

```
## # Intraclass Correlation Coefficient
##
##      Adjusted ICC: 0.210
##   Conditional ICC: 0.210
```

```
icc(mod2_1)
```

```
## # Intraclass Correlation Coefficient
##
##      Adjusted ICC: 0.673
##   Conditional ICC: 0.673
```

```
icc(mod3_1)
```

```
## # Intraclass Correlation Coefficient
##
##      Adjusted ICC: 0.687
##   Conditional ICC: 0.687
```

```r
icc(mod4_1)
```

```
## # Intraclass Correlation Coefficient
##
##      Adjusted ICC: 0.472
##   Conditional ICC: 0.472
```

```r
icc(mod5_1)
```

```
## # Intraclass Correlation Coefficient
##
##      Adjusted ICC: 0.220
##   Conditional ICC: 0.220
```

```r
icc(mod6_1)
```

```
## # Intraclass Correlation Coefficient
##
##      Adjusted ICC: 0.661
##   Conditional ICC: 0.661
```

```r
icc(mod7_1)
```

```
## # Intraclass Correlation Coefficient
##
##      Adjusted ICC: 0.188
##   Conditional ICC: 0.188
```

> Just looking at the all artifacts, and looking at each rubric separately, Rubric CritDes, InitEDA, SelMeth and VisOrg have high ICC which means high correlation between raters and ratings on the same artifact, which means that the raters are consistent with one another in how they rate. Rubric RsrchQ, InterpRes and TxtOrg all have lower ICC which means low correlation between raters and ratings on the same artifact, which means that the raters are not consistent with one another in how they rate.
>
> The ICC from the full dataset is consistent with the subset dataset ICC, so the raters's ratings do not vary across different artifacts.

RsrchQ

```r
repeated <- ratings[ratings$Repeated==1,]
raters_1_and_2_on_RsrchQ <- data.frame(r1=repeated$RsrchQ[repeated$Rater==1],r2=repeated$RsrchQ[repeate
a1=repeated$Artifact[repeated$Rater==1],
a2=repeated$Artifact[repeated$Rater==2]
)
with(raters_1_and_2_on_RsrchQ,table(r1,r2))
```

```
##    r2
## r1  1 2 3 4
##   1 0 0 0 0
##   2 1 4 3 0
##   3 1 3 1 0
##   4 0 0 0 0
```

```r
r1 <- factor(raters_1_and_2_on_RsrchQ$r1,levels=1:4)
r2 <- factor(raters_1_and_2_on_RsrchQ$r2,levels=1:4)
(t12 <- table(r1,r2))
```

```
##    r2
## r1  1 2 3 4
##   1 0 0 0 0
##   2 1 4 3 0
##   3 1 3 1 0
##   4 0 0 0 0
```

```r
5/13
```

```
## [1] 0.3846154
```

Raters group 1 and group 2 only have 38.46% of times when they give the same ratings.

```r
repeated <- ratings[ratings$Repeated==1,]
raters_2_and_3_on_RsrchQ <- data.frame(r2=repeated$RsrchQ[repeated$Rater==2],r3=repeated$RsrchQ[repeate
a1=repeated$Artifact[repeated$Rater==2],
a2=repeated$Artifact[repeated$Rater==3]
)
with(raters_2_and_3_on_RsrchQ,table(r2,r3))
```

```
##    r3
## r2  1 2 3 4
##   1 0 2 0 0
##   2 0 5 2 0
##   3 0 2 2 0
##   4 0 0 0 0
```

```r
r2 <- factor(raters_2_and_3_on_RsrchQ$r2,levels=1:4)
r3 <- factor(raters_2_and_3_on_RsrchQ$r3,levels=1:4)
(t23 <- table(r2,r3))
```

```
##    r3
## r2  1 2 3 4
##   1 0 2 0 0
##   2 0 5 2 0
##   3 0 2 2 0
##   4 0 0 0 0
```

```r
7/13
```

```
## [1] 0.5384615
```

Raters group 2 and group 3 have 53.85% of times when they give the same ratings.

```r
repeated <- ratings[ratings$Repeated==1,]
raters_1_and_3_on_RsrchQ <- data.frame(r1=repeated$RsrchQ[repeated$Rater==1],r3=repeated$RsrchQ[repeate
a1=repeated$Artifact[repeated$Rater==1],
a2=repeated$Artifact[repeated$Rater==3]
)
with(raters_1_and_3_on_RsrchQ,table(r1,r3))
```

```
##    r3
## r1  1 2 3 4
##   1 0 0 0 0
##   2 0 7 1 0
##   3 0 2 3 0
##   4 0 0 0 0
```

```
r1 <- factor(raters_1_and_3_on_RsrchQ$r1,levels=1:4)
r3 <- factor(raters_1_and_3_on_RsrchQ$r3,levels=1:4)
(t13 <- table(r1,r3))
```

```
##    r3
## r1  1 2 3 4
##   1 0 0 0 0
##   2 0 7 1 0
##   3 0 2 3 0
##   4 0 0 0 0
```

```
10/13
```

```
## [1] 0.7692308
```

Raters group 1 and group 3 have 76.92% of times when they give the same ratings.

CritDes

```
repeated <- ratings[ratings$Repeated==1,]
raters_1_and_2_on_CritDes <- data.frame(r1=repeated$CritDes[repeated$Rater==1],r2=repeated$CritDes[repea
a1=repeated$Artifact[repeated$Rater==1],
a2=repeated$Artifact[repeated$Rater==2]
)
with(raters_1_and_2_on_CritDes,table(r1,r2))
```

```
##    r2
## r1  1 2 3 4
##   1 3 2 1 0
##   2 2 3 1 0
##   3 0 0 1 0
##   4 0 0 0 0
```

```
r1 <- factor(raters_1_and_2_on_CritDes$r1,levels=1:4)
r2 <- factor(raters_1_and_2_on_CritDes$r2,levels=1:4)
(t12 <- table(r1,r2))
```

```
##    r2
## r1  1 2 3 4
##   1 3 2 1 0
##   2 2 3 1 0
##   3 0 0 1 0
##   4 0 0 0 0
```

```
7/13
```

```
## [1] 0.5384615
```

Raters group 1 and group 2 have 53.85% of times when they give the same ratings.

```
repeated <- ratings[ratings$Repeated==1,]
raters_2_and_3_on_CritDes <- data.frame(r2=repeated$CritDes[repeated$Rater==2],r3=repeated$CritDes[repea
a1=repeated$Artifact[repeated$Rater==2],
a2=repeated$Artifact[repeated$Rater==3]
)
with(raters_2_and_3_on_CritDes,table(r2,r3))
```

```
##    r3
```

```
## r2  1 2 3 4
##   1 5 0 0 0
##   2 1 3 1 0
##   3 0 2 1 0
##   4 0 0 0 0
```

```
r2 <- factor(raters_2_and_3_on_CritDes$r2,levels=1:4)
r3 <- factor(raters_2_and_3_on_CritDes$r3,levels=1:4)
(t23 <- table(r2,r3))
```

```
##    r3
## r2  1 2 3 4
##   1 5 0 0 0
##   2 1 3 1 0
##   3 0 2 1 0
##   4 0 0 0 0
```

```
9/13
```

```
## [1] 0.6923077
```

Raters group 2 and group 3 have 69.23% of times when they give the same ratings.

```
repeated <- ratings[ratings$Repeated==1,]
raters_1_and_3_on_CritDes <- data.frame(r1=repeated$CritDes[repeated$Rater==1],r3=repeated$CritDes[repea
a1=repeated$Artifact[repeated$Rater==1],
a2=repeated$Artifact[repeated$Rater==3]
)
with(raters_1_and_3_on_CritDes,table(r1,r3))
```

```
##    r3
## r1  1 2 3 4
##   1 4 2 0 0
##   2 2 3 1 0
##   3 0 0 1 0
##   4 0 0 0 0
```

```
r1 <- factor(raters_1_and_3_on_CritDes$r1,levels=1:4)
r3 <- factor(raters_1_and_3_on_CritDes$r3,levels=1:4)
(t13 <- table(r1,r3))
```

```
##    r3
## r1  1 2 3 4
##   1 4 2 0 0
##   2 2 3 1 0
##   3 0 0 1 0
##   4 0 0 0 0
```

```
8/13
```

```
## [1] 0.6153846
```

Raters group 1 and group 3 have 61.53% of times when they give the same ratings.

CritDes

```
repeated <- ratings[ratings$Repeated==1,]
raters_1_and_2_on_CritDes <- data.frame(r1=repeated$CritDes[repeated$Rater==1],r2=repeated$CritDes[repea
a1=repeated$Artifact[repeated$Rater==1],
```

```
a2=repeated$Artifact[repeated$Rater==2]
)
with(raters_1_and_2_on_CritDes,table(r1,r2))
```

```
##    r2
## r1  1 2 3 4
##   1 3 2 1 0
##   2 2 3 1 0
##   3 0 0 1 0
##   4 0 0 0 0
```

```
r1 <- factor(raters_1_and_2_on_CritDes$r1,levels=1:4)
r2 <- factor(raters_1_and_2_on_CritDes$r2,levels=1:4)
(t12 <- table(r1,r2))
```

```
##    r2
## r1  1 2 3 4
##   1 3 2 1 0
##   2 2 3 1 0
##   3 0 0 1 0
##   4 0 0 0 0
```

```
7/13
```

```
## [1] 0.5384615
```

Raters group 1 and group 2 have 53.85% of times when they give the same ratings.

```
repeated <- ratings[ratings$Repeated==1,]
raters_2_and_3_on_CritDes <- data.frame(r2=repeated$CritDes[repeated$Rater==2],r3=repeated$CritDes[repea
a1=repeated$Artifact[repeated$Rater==2],
a2=repeated$Artifact[repeated$Rater==3]
)
with(raters_2_and_3_on_CritDes,table(r2,r3))
```

```
##    r3
## r2  1 2 3 4
##   1 5 0 0 0
##   2 1 3 1 0
##   3 0 2 1 0
##   4 0 0 0 0
```

```
r2 <- factor(raters_2_and_3_on_CritDes$r2,levels=1:4)
r3 <- factor(raters_2_and_3_on_CritDes$r3,levels=1:4)
(t23 <- table(r2,r3))
```

```
##    r3
## r2  1 2 3 4
##   1 5 0 0 0
##   2 1 3 1 0
##   3 0 2 1 0
##   4 0 0 0 0
```

```
9/13
```

```
## [1] 0.6923077
```

Raters group 2 and group 3 have 69.23% of times when they give the same ratings.

```r
repeated <- ratings[ratings$Repeated==1,]
raters_1_and_3_on_CritDes <- data.frame(r1=repeated$CritDes[repeated$Rater==1],r3=repeated$CritDes[repea
a1=repeated$Artifact[repeated$Rater==1],
a2=repeated$Artifact[repeated$Rater==3]
)
with(raters_1_and_3_on_CritDes,table(r1,r3))
```

```
##    r3
## r1  1 2 3 4
##   1 4 2 0 0
##   2 2 3 1 0
##   3 0 0 1 0
##   4 0 0 0 0
```

```r
r1 <- factor(raters_1_and_3_on_CritDes$r1,levels=1:4)
r3 <- factor(raters_1_and_3_on_CritDes$r3,levels=1:4)
(t13 <- table(r1,r3))
```

```
##    r3
## r1  1 2 3 4
##   1 4 2 0 0
##   2 2 3 1 0
##   3 0 0 1 0
##   4 0 0 0 0
```

```r
8/13
```

```
## [1] 0.6153846
```

Raters group 1 and group 3 have 61.54% of times when they give the same ratings.

InitEDA

```r
repeated <- ratings[ratings$Repeated==1,]
raters_1_and_2_on_InitEDA <- data.frame(r1=repeated$InitEDA[repeated$Rater==1],r2=repeated$InitEDA[repea
a1=repeated$Artifact[repeated$Rater==1],
a2=repeated$Artifact[repeated$Rater==2]
)
with(raters_1_and_2_on_InitEDA,table(r1,r2))
```

```
##    r2
## r1  1 2 3 4
##   1 0 1 0 0
##   2 0 4 0 0
##   3 0 3 5 0
##   4 0 0 0 0
```

```r
r1 <- factor(raters_1_and_2_on_InitEDA$r1,levels=1:4)
r2 <- factor(raters_1_and_2_on_InitEDA$r2,levels=1:4)
(t12 <- table(r1,r2))
```

```
##    r2
## r1  1 2 3 4
##   1 0 1 0 0
##   2 0 4 0 0
##   3 0 3 5 0
##   4 0 0 0 0
```

```r
9/13
```

```
## [1] 0.6923077
```

Raters group 1 and group 2 have 69.23% of times when they give the same ratings.

```r
repeated <- ratings[ratings$Repeated==1,]
raters_2_and_3_on_InitEDA <- data.frame(r2=repeated$InitEDA[repeated$Rater==2],r3=repeated$InitEDA[repea
a1=repeated$Artifact[repeated$Rater==2],
a2=repeated$Artifact[repeated$Rater==3]
)
with(raters_2_and_3_on_InitEDA,table(r2,r3))
```

```
##    r3
## r2  1 2 3 4
##   1 0 0 0 0
##   2 0 8 0 0
##   3 0 2 3 0
##   4 0 0 0 0
```

```r
r2 <- factor(raters_2_and_3_on_InitEDA$r2,levels=1:4)
r3 <- factor(raters_2_and_3_on_InitEDA$r3,levels=1:4)
(t23 <- table(r2,r3))
```

```
##    r3
## r2  1 2 3 4
##   1 0 0 0 0
##   2 0 8 0 0
##   3 0 2 3 0
##   4 0 0 0 0
```

```r
11/13
```

```
## [1] 0.8461538
```

Raters group 2 and group 3 have 84.62% of times when they give the same ratings.

```r
repeated <- ratings[ratings$Repeated==1,]
raters_1_and_3_on_InitEDA <- data.frame(r1=repeated$InitEDA[repeated$Rater==1],r3=repeated$InitEDA[repea
a1=repeated$Artifact[repeated$Rater==1],
a2=repeated$Artifact[repeated$Rater==3]
)
with(raters_1_and_3_on_InitEDA,table(r1,r3))
```

```
##    r3
## r1  1 2 3 4
##   1 0 1 0 0
##   2 0 4 0 0
##   3 0 5 3 0
##   4 0 0 0 0
```

```r
r1 <- factor(raters_1_and_3_on_InitEDA$r1,levels=1:4)
r3 <- factor(raters_1_and_3_on_InitEDA$r3,levels=1:4)
(t13 <- table(r1,r3))
```

```
##    r3
## r1  1 2 3 4
##   1 0 1 0 0
##   2 0 4 0 0
```

```
##   3 0 5 3 0
##   4 0 0 0 0
```

```
7/13
```

```
## [1] 0.5384615
```

Raters group 1 and group 3 have 53.85% of times when they give the same ratings.

SelMeth

```
repeated <- ratings[ratings$Repeated==1,]
raters_1_and_2_on_SelMeth <- data.frame(r1=repeated$SelMeth[repeated$Rater==1],r2=repeated$SelMeth[repea
a1=repeated$Artifact[repeated$Rater==1],
a2=repeated$Artifact[repeated$Rater==2]
)
with(raters_1_and_2_on_SelMeth,table(r1,r2))
```

```
##    r2
## r1   1  2  3  4
##   1  0  0  0  0
##   2  1 10  0  0
##   3  0  0  2  0
##   4  0  0  0  0
```

```
r1 <- factor(raters_1_and_2_on_SelMeth$r1,levels=1:4)
r2 <- factor(raters_1_and_2_on_SelMeth$r2,levels=1:4)
(t12 <- table(r1,r2))
```

```
##    r2
## r1   1  2  3  4
##   1  0  0  0  0
##   2  1 10  0  0
##   3  0  0  2  0
##   4  0  0  0  0
```

```
12/13
```

```
## [1] 0.9230769
```

Raters group 1 and group 2 have 92.31% of times when they give the same ratings.

```
repeated <- ratings[ratings$Repeated==1,]
raters_2_and_3_on_SelMeth <- data.frame(r2=repeated$SelMeth[repeated$Rater==2],r3=repeated$SelMeth[repea
a1=repeated$Artifact[repeated$Rater==2],
a2=repeated$Artifact[repeated$Rater==3]
)
with(raters_2_and_3_on_SelMeth,table(r2,r3))
```

```
##    r3
## r2  1 2 3 4
##   1 1 0 0 0
##   2 2 7 1 0
##   3 0 1 1 0
##   4 0 0 0 0
```

```
r2 <- factor(raters_2_and_3_on_SelMeth$r2,levels=1:4)
r3 <- factor(raters_2_and_3_on_SelMeth$r3,levels=1:4)
(t23 <- table(r2,r3))
```

```
##    r3
## r2  1 2 3 4
##   1 1 0 0 0
##   2 2 7 1 0
##   3 0 1 1 0
##   4 0 0 0 0
```

```r
9/13
```

```
## [1] 0.6923077
```

Raters group 2 and group 3 have 69.23% of times when they give the same ratings.

```r
repeated <- ratings[ratings$Repeated==1,]
raters_1_and_3_on_SelMeth <- data.frame(r1=repeated$SelMeth[repeated$Rater==1],r3=repeated$SelMeth[repea
a1=repeated$Artifact[repeated$Rater==1],
a2=repeated$Artifact[repeated$Rater==3]
)
with(raters_1_and_3_on_SelMeth,table(r1,r3))
```

```
##    r3
## r1  1 2 3 4
##   1 0 0 0 0
##   2 3 7 1 0
##   3 0 1 1 0
##   4 0 0 0 0
```

```r
r1 <- factor(raters_1_and_3_on_SelMeth$r1,levels=1:4)
r3 <- factor(raters_1_and_3_on_SelMeth$r3,levels=1:4)
(t13 <- table(r1,r3))
```

```
##    r3
## r1  1 2 3 4
##   1 0 0 0 0
##   2 3 7 1 0
##   3 0 1 1 0
##   4 0 0 0 0
```

```r
8/13
```

```
## [1] 0.6153846
```

Raters group 1 and group 3 have 61.54% of times when they give the same ratings.

InterpRes

```r
repeated <- ratings[ratings$Repeated==1,]
raters_1_and_2_on_InterpRes <- data.frame(r1=repeated$InterpRes[repeated$Rater==1],r2=repeated$InterpRes
a1=repeated$Artifact[repeated$Rater==1],
a2=repeated$Artifact[repeated$Rater==2]
)
with(raters_1_and_2_on_InterpRes,table(r1,r2))
```

```
##    r2
## r1  1 2 3 4
##   1 0 0 0 0
##   2 0 3 1 1
```

```
##   3 0 3 5 0
##   4 0 0 0 0
```

```
r1 <- factor(raters_1_and_2_on_InterpRes$r1,levels=1:4)
r2 <- factor(raters_1_and_2_on_InterpRes$r2,levels=1:4)
(t12 <- table(r1,r2))
```

```
##    r2
## r1  1 2 3 4
##   1 0 0 0 0
##   2 0 3 1 1
##   3 0 3 5 0
##   4 0 0 0 0
```

```
8/13
```

```
## [1] 0.6153846
```

Raters group 1 and group 2 have 61.54% of times when they give the same ratings.

```
repeated <- ratings[ratings$Repeated==1,]
raters_2_and_3_on_InterpRes <- data.frame(r2=repeated$InterpRes[repeated$Rater==2],r3=repeated$InterpRes
a1=repeated$Artifact[repeated$Rater==2],
a2=repeated$Artifact[repeated$Rater==3]
)
with(raters_2_and_3_on_InterpRes,table(r2,r3))
```

```
##    r3
## r2  1 2 3 4
##   1 0 0 0 0
##   2 1 4 1 0
##   3 0 2 4 0
##   4 0 1 0 0
```

```
r2 <- factor(raters_2_and_3_on_InterpRes$r2,levels=1:4)
r3 <- factor(raters_2_and_3_on_InterpRes$r3,levels=1:4)
(t23 <- table(r2,r3))
```

```
##    r3
## r2  1 2 3 4
##   1 0 0 0 0
##   2 1 4 1 0
##   3 0 2 4 0
##   4 0 1 0 0
```

```
8/13
```

```
## [1] 0.6153846
```

Raters group 2 and group 3 have 61.54% of times when they give the same ratings.

```
repeated <- ratings[ratings$Repeated==1,]
raters_1_and_3_on_InterpRes <- data.frame(r1=repeated$InterpRes[repeated$Rater==1],r3=repeated$InterpRes
a1=repeated$Artifact[repeated$Rater==1],
a2=repeated$Artifact[repeated$Rater==3]
)
with(raters_1_and_3_on_InterpRes,table(r1,r3))
```

```
##    r3
## r1  1 2 3 4
```

```
##   1 0 0 0 0
##   2 1 3 1 0
##   3 0 4 4 0
##   4 0 0 0 0
```

```r
r1 <- factor(raters_1_and_3_on_InterpRes$r1,levels=1:4)
r3 <- factor(raters_1_and_3_on_InterpRes$r3,levels=1:4)
(t13 <- table(r1,r3))
```

```
##    r3
## r1  1 2 3 4
##   1 0 0 0 0
##   2 1 3 1 0
##   3 0 4 4 0
##   4 0 0 0 0
```

```r
7/13
```

```
## [1] 0.5384615
```

Raters group 1 and group 3 have 53.85% of times when they give the same ratings.

VisOrg

```r
repeated <- ratings[ratings$Repeated==1,]
raters_1_and_2_on_VisOrg <- data.frame(r1=repeated$VisOrg[repeated$Rater==1],r2=repeated$VisOrg[repeate
a1=repeated$Artifact[repeated$Rater==1],
a2=repeated$Artifact[repeated$Rater==2]
)
with(raters_1_and_2_on_VisOrg,table(r1,r2))
```

```
##    r2
## r1  1 2 3 4
##   1 1 0 0 0
##   2 0 4 5 0
##   3 0 1 2 0
##   4 0 0 0 0
```

```r
r1 <- factor(raters_1_and_2_on_VisOrg$r1,levels=1:4)
r2 <- factor(raters_1_and_2_on_VisOrg$r2,levels=1:4)
(t12 <- table(r1,r2))
```

```
##    r2
## r1  1 2 3 4
##   1 1 0 0 0
##   2 0 4 5 0
##   3 0 1 2 0
##   4 0 0 0 0
```

```r
7/13
```

```
## [1] 0.5384615
```

Raters group 1 and group 2 have 53.85% of times when they give the same ratings.

```r
repeated <- ratings[ratings$Repeated==1,]
raters_2_and_3_on_VisOrg <- data.frame(r2=repeated$VisOrg[repeated$Rater==2],r3=repeated$VisOrg[repeate
a1=repeated$Artifact[repeated$Rater==2],
a2=repeated$Artifact[repeated$Rater==3]
```

```r
)
with(raters_2_and_3_on_VisOrg,table(r2,r3))
```

```
##    r3
## r2  1 2 3 4
##   1 1 0 0 0
##   2 0 5 0 0
##   3 0 3 4 0
##   4 0 0 0 0
```

```r
r2 <- factor(raters_2_and_3_on_VisOrg$r2,levels=1:4)
r3 <- factor(raters_2_and_3_on_VisOrg$r3,levels=1:4)
(t23 <- table(r2,r3))
```

```
##    r3
## r2  1 2 3 4
##   1 1 0 0 0
##   2 0 5 0 0
##   3 0 3 4 0
##   4 0 0 0 0
```

```r
10/13
```

```
## [1] 0.7692308
```

Raters group 2 and group 3 have 76.92% of times when they give the same ratings.

```r
repeated <- ratings[ratings$Repeated==1,]
raters_1_and_3_on_VisOrg <- data.frame(r1=repeated$VisOrg[repeated$Rater==1],r3=repeated$VisOrg[repeatec
a1=repeated$Artifact[repeated$Rater==1],
a2=repeated$Artifact[repeated$Rater==3]
)
with(raters_1_and_3_on_VisOrg,table(r1,r3))
```

```
##    r3
## r1  1 2 3 4
##   1 1 0 0 0
##   2 0 7 2 0
##   3 0 1 2 0
##   4 0 0 0 0
```

```r
r1 <- factor(raters_1_and_3_on_VisOrg$r1,levels=1:4)
r3 <- factor(raters_1_and_3_on_VisOrg$r3,levels=1:4)
(t13 <- table(r1,r3))
```

```
##    r3
## r1  1 2 3 4
##   1 1 0 0 0
##   2 0 7 2 0
##   3 0 1 2 0
##   4 0 0 0 0
```

```r
10/13
```

```
## [1] 0.7692308
```

Raters group 1 and group 3 have 76.92% of times when they give the same ratings.

TxtOrg

```
repeated <- ratings[ratings$Repeated==1,]
raters_1_and_2_on_TxtOrg <- data.frame(r1=repeated$TxtOrg[repeated$Rater==1],r2=repeated$TxtOrg[repeatec
a1=repeated$Artifact[repeated$Rater==1],
a2=repeated$Artifact[repeated$Rater==2]
)
with(raters_1_and_2_on_VisOrg,table(r1,r2))
```

```
##    r2
## r1  1 2 3 4
##   1 1 0 0 0
##   2 0 4 5 0
##   3 0 1 2 0
##   4 0 0 0 0
```

```
r1 <- factor(raters_1_and_2_on_TxtOrg$r1,levels=1:4)
r2 <- factor(raters_1_and_2_on_TxtOrg$r2,levels=1:4)
(t12 <- table(r1,r2))
```

```
##    r2
## r1  1 2 3 4
##   1 0 0 0 0
##   2 0 2 2 0
##   3 0 1 7 0
##   4 1 0 0 0
```

```
9/13
```

```
## [1] 0.6923077
```

Raters group 1 and group 2 have 69.23% of times when they give the same ratings.

```
repeated <- ratings[ratings$Repeated==1,]
raters_2_and_3_on_TxtOrg <- data.frame(r2=repeated$TxtOrg[repeated$Rater==2],r3=repeated$TxtOrg[repeatec
a1=repeated$Artifact[repeated$Rater==2],
a2=repeated$Artifact[repeated$Rater==3]
)
with(raters_2_and_3_on_TxtOrg,table(r2,r3))
```

```
##    r3
## r2  1 2 3 4
##   1 0 1 0 0
##   2 1 0 2 0
##   3 0 2 7 0
##   4 0 0 0 0
```

```
r2 <- factor(raters_2_and_3_on_TxtOrg$r2,levels=1:4)
r3 <- factor(raters_2_and_3_on_TxtOrg$r3,levels=1:4)
(t23 <- table(r2,r3))
```

```
##    r3
## r2  1 2 3 4
##   1 0 1 0 0
##   2 1 0 2 0
##   3 0 2 7 0
##   4 0 0 0 0
```

```r
7/13
```

```
## [1] 0.5384615
```

Raters group 2 and group 3 have 53.85% of times when they give the same ratings.

```r
repeated <- ratings[ratings$Repeated==1,]
raters_1_and_3_on_TxtOrg <- data.frame(r1=repeated$TxtOrg[repeated$Rater==1],r3=repeated$TxtOrg[repeatec
a1=repeated$Artifact[repeated$Rater==1],
a2=repeated$Artifact[repeated$Rater==3]
)
with(raters_1_and_3_on_TxtOrg,table(r1,r3))
```

```
##     r3
## r1   1 2 3 4
##    1 0 0 0 0
##    2 1 1 2 0
##    3 0 1 7 0
##    4 0 1 0 0
```

```r
r1 <- factor(raters_1_and_3_on_TxtOrg$r1,levels=1:4)
r3 <- factor(raters_1_and_3_on_TxtOrg$r3,levels=1:4)
(t13 <- table(r1,r3))
```

```
##     r3
## r1   1 2 3 4
##    1 0 0 0 0
##    2 1 1 2 0
##    3 0 1 7 0
##    4 0 1 0 0
```

```r
8/13
```

```
## [1] 0.6153846
```

Raters group 1 and group 3 have 61.53% of times when they give the same ratings.

```r
fullmod1 <- lmer(as.numeric(Rating) ~ 1 + (1|Artifact), data=RsrchQ.full)
fullmod2 <- lmer(as.numeric(Rating) ~ 1 + (1|Artifact), data=CritDes.full)
fullmod3 <- lmer(as.numeric(Rating) ~ 1 + (1|Artifact), data=InitEDA.full)
fullmod4 <- lmer(as.numeric(Rating) ~ 1 + (1|Artifact), data=SelMeth.full)
fullmod5 <- lmer(as.numeric(Rating) ~ 1 + (1|Artifact), data=InterpRes.full)
fullmod6 <- lmer(as.numeric(Rating) ~ 1 + (1|Artifact), data=VisOrg.full)
fullmod7 <- lmer(as.numeric(Rating) ~ 1 + (1|Artifact), data=TxtOrg.full)
```

```r
all.mod <- lmer(as.numeric(Rating) ~ 1 + Sex + Semester + (1|Artifact), data=tall)
icc(all.mod)
```

```
## # Intraclass Correlation Coefficient
##
##      Adjusted ICC: 0.260
##   Conditional ICC: 0.257
```

Part 3

RsrchQ

```
library(LMERConvenienceFunctions)
library(RLRsim)

tmp <- lmer(as.numeric(Rating) ~ -1 + as.factor(Rater) +
              Semester + Sex + (1|Artifact),
            data=tall.13[tall.13$Rubric=="RsrchQ",],REML=FALSE)

tmp.back_elim <- fitLMER.fnc(tmp,set.REML.FALSE = TRUE,log.file.name = FALSE)
```

```
## ======================================================
## ===              backfitting fixed effects          ===
## ======================================================
## processing model terms of interaction level 1
##   iteration 1
##     p-value for term "Semester" = 0.7355 >= 0.05
##     not part of higher-order interaction
##     removing term
##   iteration 2
##     p-value for term "Sex" = 0.279 >= 0.05
##     not part of higher-order interaction
##     removing term
## pruning random effects structure ...
##   nothing to prune
## ======================================================
## ===            forwardfitting random effects        ===
## ======================================================
##  ===       random slopes       ===
## ======================================================
## ===            re-backfitting fixed effects         ===
## ======================================================
## processing model terms of interaction level 1
##   all terms of interaction level 1 significant
## resetting REML to TRUE
## pruning random effects structure ...
##   nothing to prune
```

```
formula(tmp.back_elim)
```

```
## as.numeric(Rating) ~ as.factor(Rater) + (1 | Artifact) - 1
```

```
tmp.int_only <- update(tmp.back_elim, . ~ . + 1 - as.factor(Rater))

anova(tmp.int_only,tmp.back_elim)
```

```
## Data: tall.13[tall.13$Rubric == "RsrchQ", ]
## Models:
## tmp.int_only: as.numeric(Rating) ~ (1 | Artifact)
## tmp.back_elim: as.numeric(Rating) ~ as.factor(Rater) + (1 | Artifact) - 1
##               npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## tmp.int_only     3 69.457 74.447 -31.728   63.457
## tmp.back_elim    5 72.018 80.335 -31.009   62.018 1.4391  2      0.487
```

```
anova(tmp.int_only,tmp.back_elim)$"Pr(>Chisq)"[2]
```

```
## [1] 0.4869707
```

```r
Rubric.names <- sort(unique(tall$Rubric))

model.formula.13 <- as.list(rep(NA,7))
names(model.formula.13) <- Rubric.names



for (i in Rubric.names) {

  ## fit each base model
  rubric.data <- tall.13[tall.13$Rubric==i,]
  tmp <- lmer(as.numeric(Rating) ~ -1 + as.factor(Rater) +
               Semester + Sex + (1|Artifact),
             data=rubric.data,REML=FALSE)

  ## do backwards elimination
  tmp.back_elim <- fitLMER.fnc(tmp,set.REML.FALSE = TRUE,log.file.name = FALSE)

  ## check to see if the raters are significantly different from one another
  tmp.single_intercept <- update(tmp.back_elim, . ~ . + 1 - as.factor(Rater))
  pval <- anova(tmp.single_intercept,tmp.back_elim)$"Pr(>Chisq)"[2]

  ## choose the best model
  if (pval<=0.05) {
    tmp_final <- tmp.back_elim
  } else {
    tmp_final <- tmp.single_intercept
  }

  ## and add to list...
  model.formula.13[[i]] <- formula(tmp_final)

}
```

```
## ======================================================
## ===              backfitting fixed effects         ===
## ======================================================
## processing model terms of interaction level 1
##   iteration 1
##     p-value for term "Sex" = 0.2229 >= 0.05
##     not part of higher-order interaction
##     removing term
##   iteration 2
##     p-value for term "Semester" = 0.1826 >= 0.05
##     not part of higher-order interaction
##     removing term
## pruning random effects structure ...
##   nothing to prune
## ======================================================
## ===            forwardfitting random effects       ===
## ======================================================
##  ===       random slopes       ===
## ======================================================
```

```
## ===            re-backfitting fixed effects        ===
## ========================================================
## processing model terms of interaction level 1
##   all terms of interaction level 1 significant
## resetting REML to TRUE
## pruning random effects structure ...
##   nothing to prune
## ========================================================
## ===              backfitting fixed effects         ===
## ========================================================
## processing model terms of interaction level 1
##   iteration 1
##     p-value for term "Semester" = 0.8137 >= 0.05
##     not part of higher-order interaction
##     removing term
##   iteration 2
##     p-value for term "Sex" = 0.6429 >= 0.05
##     not part of higher-order interaction
##     removing term
## pruning random effects structure ...
##   nothing to prune
## ========================================================
## ===            forwardfitting random effects       ===
## ========================================================
##  ===       random slopes       ===
## ========================================================
## ===            re-backfitting fixed effects        ===
## ========================================================
## processing model terms of interaction level 1
##   all terms of interaction level 1 significant
## resetting REML to TRUE
## pruning random effects structure ...
##   nothing to prune
## ========================================================
## ===              backfitting fixed effects         ===
## ========================================================
## processing model terms of interaction level 1
##   iteration 1
##     p-value for term "Semester" = 0.8294 >= 0.05
##     not part of higher-order interaction
##     removing term
##   iteration 2
##     p-value for term "Sex" = 0.2947 >= 0.05
##     not part of higher-order interaction
##     removing term
## pruning random effects structure ...
##   nothing to prune
## ========================================================
## ===            forwardfitting random effects       ===
## ========================================================
##  ===       random slopes       ===
## ========================================================
## ===            re-backfitting fixed effects        ===
## ========================================================
```

```
## processing model terms of interaction level 1
##   all terms of interaction level 1 significant
## resetting REML to TRUE
## pruning random effects structure ...
##   nothing to prune
## ======================================================
## ===              backfitting fixed effects         ===
## ======================================================
## processing model terms of interaction level 1
##   iteration 1
##     p-value for term "Semester" = 0.7355 >= 0.05
##     not part of higher-order interaction
##     removing term
##   iteration 2
##     p-value for term "Sex" = 0.279 >= 0.05
##     not part of higher-order interaction
##     removing term
## pruning random effects structure ...
##   nothing to prune
## ======================================================
## ===            forwardfitting random effects       ===
## ======================================================
##  ===       random slopes       ===
## ======================================================
## ===            re-backfitting fixed effects        ===
## ======================================================
## processing model terms of interaction level 1
##   all terms of interaction level 1 significant
## resetting REML to TRUE
## pruning random effects structure ...
##   nothing to prune
## ======================================================
## ===              backfitting fixed effects         ===
## ======================================================
## processing model terms of interaction level 1
##   iteration 1
##     p-value for term "Sex" = 0.9383 >= 0.05
##     not part of higher-order interaction
##     removing term
##   iteration 2
##     p-value for term "Semester" = 0.4287 >= 0.05
##     not part of higher-order interaction
##     removing term
## pruning random effects structure ...
##   nothing to prune
## ======================================================
## ===            forwardfitting random effects       ===
## ======================================================
##  ===       random slopes       ===
## ======================================================
## ===            re-backfitting fixed effects        ===
## ======================================================
## processing model terms of interaction level 1
##   all terms of interaction level 1 significant
```

```
## resetting REML to TRUE
## pruning random effects structure ...
##   nothing to prune
## =========================================================
## ===              backfitting fixed effects          ===
## =========================================================
## processing model terms of interaction level 1
##   iteration 1
##     p-value for term "Semester" = 0.5358 >= 0.05
##     not part of higher-order interaction
##     removing term
##   iteration 2
##     p-value for term "Sex" = 0.1319 >= 0.05
##     not part of higher-order interaction
##     removing term
## pruning random effects structure ...
##   nothing to prune
## =========================================================
## ===            forwardfitting random effects        ===
## =========================================================
##  ===       random slopes       ===
## =========================================================
## ===            re-backfitting fixed effects         ===
## =========================================================
## processing model terms of interaction level 1
##   all terms of interaction level 1 significant
## resetting REML to TRUE
## pruning random effects structure ...
##   nothing to prune
## =========================================================
## ===              backfitting fixed effects          ===
## =========================================================
## processing model terms of interaction level 1
##   iteration 1
##     p-value for term "Semester" = 0.1922 >= 0.05
##     not part of higher-order interaction
##     removing term
##   iteration 2
##     p-value for term "Sex" = 0.1078 >= 0.05
##     not part of higher-order interaction
##     removing term
## pruning random effects structure ...
##   nothing to prune
## =========================================================
## ===            forwardfitting random effects        ===
## =========================================================
##  ===       random slopes       ===
## =========================================================
## ===            re-backfitting fixed effects         ===
## =========================================================
## processing model terms of interaction level 1
##   all terms of interaction level 1 significant
## resetting REML to TRUE
## pruning random effects structure ...
```

```
##   nothing to prune
```

```r
## see what "final models" we got...
model.formula.13
```

```
## $CritDes
## as.numeric(Rating) ~ (1 | Artifact)
##
## $InitEDA
## as.numeric(Rating) ~ (1 | Artifact)
##
## $InterpRes
## as.numeric(Rating) ~ (1 | Artifact)
##
## $RsrchQ
## as.numeric(Rating) ~ (1 | Artifact)
##
## $SelMeth
## as.numeric(Rating) ~ (1 | Artifact)
##
## $TxtOrg
## as.numeric(Rating) ~ (1 | Artifact)
##
## $VisOrg
## as.numeric(Rating) ~ (1 | Artifact)
```

It looks like we don't need to add any fixed effects or interactions to the models for each rubric, using only the data reduced to the 13 common rubrics. ANOVA test shows that intercept-only models are better than models with fixed effects chosen after back elimination.

```r
m1 <- lmer(as.numeric(Rating) ~ (1 | Artifact), data = tall.13[tall.13$Rubric=="CritDes",])
summary(m1)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: as.numeric(Rating) ~ (1 | Artifact)
##    Data: tall.13[tall.13$Rubric == "CritDes", ]
##
## REML criterion at convergence: 75.1
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.9647 -0.4386 -0.2978  0.5318  2.1987
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Artifact (Intercept) 0.3091   0.5560
##  Residual             0.2308   0.4804
## Number of obs: 39, groups:  Artifact, 13
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)   1.7179     0.1723   9.969
```

```r
ranef(m1)
```

```
## $Artifact
##     (Intercept)
```

```
## O1  -0.30797616
## O10 -0.30797616
## O11 -0.57488883
## O12 -0.04106349
## O13  0.22584918
## O2  -0.30797616
## O3   0.49276186
## O4   0.22584918
## O5   1.02658720
## O6  -0.57488883
## O7   0.49276186
## O8   0.22584918
## O9  -0.57488883
##
## with conditional variances for "Artifact"
```

The result is different from the result from ANOVA test as all intercept-only models are reccommended with no fixed effect.

rubric specific model

```
Rubric.names <- sort(unique(tall$Rubric))

tall[c(161,684),]
```

```
##       X Rater Artifact Repeated Semester Sex  Rubric Rating
## 162 162     2       46        0      S19   F CritDes      2
## 686 686     1      102        0      F19   M  VisOrg      2
```

```
tall.nonmissing <- tall[-c(161,684),]


tall.nonmissing[tall.nonmissing$Sex=="--",]
```

```
##       X Rater Artifact Repeated Semester Sex    Rubric Rating
## 5     5     3        5        0      F19  --    RsrchQ      3
## 122 122     3        5        0      F19  --   CritDes      3
## 239 239     3        5        0      F19  --   InitEDA      3
## 356 356     3        5        0      F19  --   SelMeth      3
## 473 473     3        5        0      F19  -- InterpRes      3
## 590 590     3        5        0      F19  --    VisOrg      3
## 707 707     3        5        0      F19  --    TxtOrg      3
```

```
tall.nonmissing <- tall.nonmissing[tall.nonmissing$Sex!="--",]

model.formula.alldata <- as.list(rep(NA,7))
names(model.formula.alldata) <- Rubric.names

for (i in Rubric.names) {

  rubric.data <- tall.nonmissing[tall.nonmissing$Rubric==i,]
  tmp <- lmer(as.numeric(Rating) ~ -1 + as.factor(Rater) +
                Semester + Sex + (1|Artifact),
              data=rubric.data,REML=FALSE)

  tmp.back_elim <- fitLMER.fnc(tmp,set.REML.FALSE = TRUE,log.file.name = FALSE)
```

```r
  tmp.single_intercept <- update(tmp.back_elim, . ~ . + 1 - as.factor(Rater))
  pval <- anova(tmp.single_intercept,tmp.back_elim)$"Pr(>Chisq)"[2]

  if (pval<=0.05) {
    tmp_final <- tmp.back_elim
  } else {
    tmp_final <- tmp.single_intercept
  }

  model.formula.alldata[[i]] <- formula(tmp_final)

}
```

```
## Warning in fitLMER.fnc(tmp, set.REML.FALSE = TRUE, log.file.name = FALSE): Argument "ran.effects" is
## TRUE

## ======================================================
## ===              backfitting fixed effects          ===
## ======================================================
## processing model terms of interaction level 1
##   iteration 1
##     p-value for term "Semester" = 0.7308 >= 0.05
##     not part of higher-order interaction
##     removing term
##   iteration 2
##     p-value for term "Sex" = 0.5393 >= 0.05
##     not part of higher-order interaction
##     removing term
## pruning random effects structure ...
##   nothing to prune
## ======================================================
## ===            forwardfitting random effects        ===
## ======================================================
##  ===       random slopes       ===
## ======================================================
## ===            re-backfitting fixed effects         ===
## ======================================================
## processing model terms of interaction level 1
##   all terms of interaction level 1 significant
## resetting REML to TRUE
## pruning random effects structure ...
##   nothing to prune

## refitting model(s) with ML (instead of REML)

## Warning in fitLMER.fnc(tmp, set.REML.FALSE = TRUE, log.file.name = FALSE): Argument "ran.effects" is
## TRUE

## ======================================================
## ===              backfitting fixed effects          ===
## ======================================================
## processing model terms of interaction level 1
##   iteration 1
##     p-value for term "Semester" = 0.8802 >= 0.05
##     not part of higher-order interaction
```

```
##     removing term
##   iteration 2
##     p-value for term "Sex" = 0.7402 >= 0.05
##     not part of higher-order interaction
##     removing term
## pruning random effects structure ...
##   nothing to prune
## ======================================================
## ===            forwardfitting random effects       ===
## ======================================================
##  ===       random slopes       ===
## ======================================================
## ===            re-backfitting fixed effects        ===
## ======================================================
## processing model terms of interaction level 1
##   all terms of interaction level 1 significant
## resetting REML to TRUE
## pruning random effects structure ...
##   nothing to prune

## refitting model(s) with ML (instead of REML)

## Warning in fitLMER.fnc(tmp, set.REML.FALSE = TRUE, log.file.name = FALSE): Argument "ran.effects" is
## TRUE

## ======================================================
## ===              backfitting fixed effects         ===
## ======================================================
## processing model terms of interaction level 1
##   iteration 1
##     p-value for term "Sex" = 0.608 >= 0.05
##     not part of higher-order interaction
##     removing term
##   iteration 2
##     p-value for term "Semester" = 0.5312 >= 0.05
##     not part of higher-order interaction
##     removing term
## pruning random effects structure ...
##   nothing to prune
## ======================================================
## ===            forwardfitting random effects       ===
## ======================================================
##  ===       random slopes       ===
## ======================================================
## ===            re-backfitting fixed effects        ===
## ======================================================
## processing model terms of interaction level 1
##   all terms of interaction level 1 significant
## resetting REML to TRUE
## pruning random effects structure ...
##   nothing to prune

## refitting model(s) with ML (instead of REML)

## Warning in fitLMER.fnc(tmp, set.REML.FALSE = TRUE, log.file.name = FALSE): Argument "ran.effects" is
## TRUE
```

```
## =========================================================
## ===              backfitting fixed effects           ===
## =========================================================
## processing model terms of interaction level 1
##   iteration 1
##     p-value for term "Sex" = 0.6166 >= 0.05
##     not part of higher-order interaction
##     removing term
##   iteration 2
##     p-value for term "Semester" = 0.3987 >= 0.05
##     not part of higher-order interaction
##     removing term
## pruning random effects structure ...
##   nothing to prune
## =========================================================
## ===            forwardfitting random effects         ===
## =========================================================
##  ===       random slopes       ===
## =========================================================
## ===            re-backfitting fixed effects          ===
## =========================================================
## processing model terms of interaction level 1
##   all terms of interaction level 1 significant
## resetting REML to TRUE
## pruning random effects structure ...
##   nothing to prune

## refitting model(s) with ML (instead of REML)

## Warning in fitLMER.fnc(tmp, set.REML.FALSE = TRUE, log.file.name = FALSE): Argument "ran.effects" is
## TRUE

## =========================================================
## ===              backfitting fixed effects           ===
## =========================================================
## processing model terms of interaction level 1
##   iteration 1
##     p-value for term "Sex" = 0.1935 >= 0.05
##     not part of higher-order interaction
##     removing term
## pruning random effects structure ...
##   nothing to prune
## =========================================================
## ===            forwardfitting random effects         ===
## =========================================================
##  ===       random slopes       ===
## =========================================================
## ===            re-backfitting fixed effects          ===
## =========================================================
## processing model terms of interaction level 1
##   all terms of interaction level 1 significant
## resetting REML to TRUE
## pruning random effects structure ...
##   nothing to prune

## refitting model(s) with ML (instead of REML)
```

```
## Warning in fitLMER.fnc(tmp, set.REML.FALSE = TRUE, log.file.name = FALSE): Argument "ran.effects" is
## TRUE

## ========================================================
## ===              backfitting fixed effects         ===
## ========================================================
## processing model terms of interaction level 1
##   iteration 1
##     p-value for term "Sex" = 0.5041 >= 0.05
##     not part of higher-order interaction
##     removing term
##   iteration 2
##     p-value for term "Semester" = 0.205 >= 0.05
##     not part of higher-order interaction
##     removing term
## pruning random effects structure ...
##   nothing to prune
## ========================================================
## ===            forwardfitting random effects       ===
## ========================================================
##  ===       random slopes       ===
## ========================================================
## ===            re-backfitting fixed effects        ===
## ========================================================
## processing model terms of interaction level 1
##   all terms of interaction level 1 significant
## resetting REML to TRUE
## pruning random effects structure ...
##   nothing to prune

## refitting model(s) with ML (instead of REML)

## Warning in fitLMER.fnc(tmp, set.REML.FALSE = TRUE, log.file.name = FALSE): Argument "ran.effects" is
## TRUE

## ========================================================
## ===              backfitting fixed effects         ===
## ========================================================
## processing model terms of interaction level 1
##   iteration 1
##     p-value for term "Semester" = 0.2007 >= 0.05
##     not part of higher-order interaction
##     removing term
##   iteration 2
##     p-value for term "Sex" = 0.3884 >= 0.05
##     not part of higher-order interaction
##     removing term
## pruning random effects structure ...
##   nothing to prune
## ========================================================
## ===            forwardfitting random effects       ===
## ========================================================
##  ===       random slopes       ===
## ========================================================
## ===            re-backfitting fixed effects        ===
## ========================================================
```

```
## processing model terms of interaction level 1
##   all terms of interaction level 1 significant
## resetting REML to TRUE
## pruning random effects structure ...
##   nothing to prune
## refitting model(s) with ML (instead of REML)
```

```
model.formula.alldata
```

```
## $CritDes
## as.numeric(Rating) ~ as.factor(Rater) + (1 | Artifact) - 1
##
## $InitEDA
## as.numeric(Rating) ~ (1 | Artifact)
##
## $InterpRes
## as.numeric(Rating) ~ as.factor(Rater) + (1 | Artifact) - 1
##
## $RsrchQ
## as.numeric(Rating) ~ (1 | Artifact)
##
## $SelMeth
## as.numeric(Rating) ~ as.factor(Rater) + Semester + (1 | Artifact) -
##     1
##
## $TxtOrg
## as.numeric(Rating) ~ (1 | Artifact)
##
## $VisOrg
## as.numeric(Rating) ~ as.factor(Rater) + (1 | Artifact) - 1
```

For InitEDA, RsrchQ, and TxtOrg, adding fixed effects does not improve the fit of the model. However, adding Rater improves the fit for CritDes, InterpRes, and VisOrg, and adding Rater, adding Semester improves the fit for SelMeth.

```
fla <- formula(model.formula.alldata[["SelMeth"]])
tmp <- lmer(fla,data=tall.nonmissing[tall.nonmissing$Rubric=="SelMeth",])
round(summary(tmp)$coef,2)
```

```
##                     Estimate Std. Error t value
## as.factor(Rater)1     2.25       0.08   29.99
## as.factor(Rater)2     2.23       0.07   29.99
## as.factor(Rater)3     2.03       0.08   27.03
## SemesterS19          -0.36       0.10   -3.66
```

```
tmp.single_intercept <- update(tmp, . ~ . + 1 - as.factor(Rater))
anova(tmp.single_intercept,tmp)
```

```
## refitting model(s) with ML (instead of REML)
## Data: tall.nonmissing[tall.nonmissing$Rubric == "SelMeth", ]
## Models:
## tmp.single_intercept: as.numeric(Rating) ~ Semester + (1 | Artifact)
## tmp: as.numeric(Rating) ~ as.factor(Rater) + Semester + (1 | Artifact) - 1
##                      npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## tmp.single_intercept    4 145.07 156.08 -68.534   137.07
## tmp                     6 142.05 158.58 -65.027   130.05 7.0146  2    0.02998 *
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
tmp.fixed_interactions <- update(tmp, . ~ . + as.factor(Rater)*Semester - Semester)

anova(tmp,tmp.fixed_interactions)
```

```
## refitting model(s) with ML (instead of REML)

## Data: tall.nonmissing[tall.nonmissing$Rubric == "SelMeth", ]
## Models:
## tmp: as.numeric(Rating) ~ as.factor(Rater) + Semester + (1 | Artifact) - 1
## tmp.fixed_interactions: as.numeric(Rating) ~ as.factor(Rater) + (1 | Artifact) + as.factor(Rater):Ser
##                        npar    AIC    BIC  logLik deviance Chisq Df Pr(>Chisq)
## tmp                       6 142.05 158.58 -65.027   130.05
## tmp.fixed_interactions    8 143.46 165.49 -63.731   127.46 2.592  2     0.2736
```

```r
m0 <- tmp                                       ## Null hypothesis
mA <- update(m0, . ~ . + (Semester|Artifact))   ## Alternative hypotheses
```

```
## Error: number of observations (=116) <= number of random effects (=180) for term (Semester | Artifac
```

```r
m  <- update(mA, . ~ . - (1|Artifact))          ## Model with only the new R.E.
```

```
## Error in h(simpleError(msg, call)): error in evaluating the argument 'object' in selecting a method f
```

```r
exactRLRT(m0=m0,mA=mA,m=m)
```

```
## Error in exactRLRT(m0 = m0, mA = mA, m = m): object 'm' not found
```

```r
m0 <- tmp                                       ## Null hypothesis
mA <- update(m0, . ~ . + (as.factor(Rater)|Artifact))   ## Alternative hypotheses
```

```
## Error: number of observations (=116) <= number of random effects (=270) for term (as.factor(Rater) |
```

```r
m  <- update(mA, . ~ . - (1|Artifact))          ## Model with only the new R.E.
```

```
## Error in h(simpleError(msg, call)): error in evaluating the argument 'object' in selecting a method f
```

```r
exactRLRT(m0=m0,mA=mA,m=m)
```

```
## Error in exactRLRT(m0 = m0, mA = mA, m = m): object 'm' not found
```

```r
summary(tmp)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: as.numeric(Rating) ~ as.factor(Rater) + Semester + (1 | Artifact) -
##     1
##    Data: tall.nonmissing[tall.nonmissing$Rubric == "SelMeth", ]
##
## REML criterion at convergence: 143.6
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.0480 -0.3923 -0.0551  0.2674  2.5827
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Artifact (Intercept) 0.08973  0.2996
##  Residual             0.10842  0.3293
## Number of obs: 116, groups:  Artifact, 90
```

```
##
## Fixed effects:
##                    Estimate Std. Error t value
## as.factor(Rater)1  2.25037    0.07503  29.992
## as.factor(Rater)2  2.22653    0.07424  29.991
## as.factor(Rater)3  2.03316    0.07521  27.033
## SemesterS19       -0.35860    0.09796  -3.661
##
## Correlation of Fixed Effects:
##              a.(R)1 a.(R)2 a.(R)3
## as.fctr(R)2  0.285
## as.fctr(R)3  0.287  0.280
## SemesterS19 -0.413 -0.391 -0.394
```

```
head(ranef(tmp),10)
```

```
## $Artifact
##     (Intercept)
## 100  0.33946601
## 101  0.04901108
## 102 -0.11338077
## 103  0.33946601
## 104 -0.11338077
## 105 -0.11338077
## 106 -0.11338077
## 107 -0.11338077
## 111  0.04901108
## 112 -0.11338077
## 113  0.04901108
## 114  0.04901108
## 115  0.04901108
## 116 -0.11338077
## 117 -0.11338077
## 118 -0.11338077
## 13  -0.46786343
## 15  -0.01501665
## 16  -0.01501665
## 17  -0.30547158
## 21   0.14737520
## 22  -0.01501665
## 23  -0.30547158
## 24  -0.01501665
## 25  -0.30547158
## 26   0.43783013
## 27  -0.01501665
## 28  -0.30547158
## 32  -0.01501665
## 33   0.43783013
## 34   0.43783013
## 35  -0.01501665
## 36  -0.01501665
## 37  -0.01501665
## 38  -0.01501665
## 39   0.14737520
## 40  -0.01501665
```

```
## 45   0.05980677
## 46   0.05980677
## 47  -0.39304001
## 48   0.35026170
## 49  -0.10258508
## 53   0.35026170
## 54  -0.10258508
## 55  -0.10258508
## 56  -0.10258508
## 57  -0.10258508
## 6   -0.01501665
## 61  -0.10258508
## 62   0.05980677
## 63   0.05980677
## 64  -0.10258508
## 65  -0.10258508
## 66   0.05980677
## 67  -0.10258508
## 68   0.05980677
## 7   -0.01501665
## 72   0.05980677
## 73  -0.10258508
## 74   0.35026170
## 75  -0.10258508
## 76  -0.10258508
## 77  -0.10258508
## 78   0.35026170
## 79   0.35026170
## 8    0.14737520
## 84   0.04901108
## 85  -0.11338077
## 86   0.04901108
## 87  -0.11338077
## 88   0.04901108
## 9    0.14737520
## 92  -0.11338077
## 93   0.04901108
## 94  -0.11338077
## 95   0.33946601
## 96  -0.11338077
## O1  -0.35883486
## O10 -0.12120653
## O11  0.13443559
## O12 -0.12120653
## O13 -0.12120653
## O2  -0.59646319
## O3  -0.12120653
## O4   0.59167847
## O5   0.35405014
## O6  -0.12120653
## O7   0.11642180
## O8  -0.10319274
## O9   0.13443559
```

CritDes

```
fla <- formula(model.formula.alldata[["CritDes"]])
tmp <- lmer(fla,data=tall.nonmissing[tall.nonmissing$Rubric=="CritDes",])
round(summary(tmp)$coef,2)
```

```
##                    Estimate Std. Error t value
## as.factor(Rater)1     1.69       0.12   13.92
## as.factor(Rater)2     2.12       0.12   17.09
## as.factor(Rater)3     1.89       0.12   15.44
```

```
tmp.single_intercept <- update(tmp, . ~ . + 1 - as.factor(Rater))
anova(tmp.single_intercept,tmp)
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: tall.nonmissing[tall.nonmissing$Rubric == "CritDes", ]
## Models:
## tmp.single_intercept: as.numeric(Rating) ~ (1 | Artifact)
## tmp: as.numeric(Rating) ~ as.factor(Rater) + (1 | Artifact) - 1
##                      npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## tmp.single_intercept    3 276.14 284.35 -135.07   270.14
## tmp                     5 272.17 285.85 -131.08   262.17 7.9767  2    0.01853 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
m0 <- tmp                                       ## Null hypothesis
mA <- update(m0, . ~ . + (as.factor(Rater)|Artifact))   ## Alternative hypotheses
```

```
## Error: number of observations (=114) <= number of random effects (=264) for term (as.factor(Rater) |
```

```
m  <- update(mA, . ~ . - (1|Artifact))          ## Model with only the new R.E.
```

```
## Error in h(simpleError(msg, call)): error in evaluating the argument 'object' in selecting a method
```

```
exactRLRT(m0=m0,mA=mA,m=m)
```

```
## Error in exactRLRT(m0 = m0, mA = mA, m = m): object 'm' not found
```

```
summary(tmp)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: as.numeric(Rating) ~ as.factor(Rater) + (1 | Artifact) - 1
##    Data: tall.nonmissing[tall.nonmissing$Rubric == "CritDes", ]
##
## REML criterion at convergence: 269.5
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.55416 -0.49643 -0.08305  0.63897  1.60548
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Artifact (Intercept) 0.4416   0.6645
##  Residual             0.2477   0.4977
## Number of obs: 114, groups:  Artifact, 88
##
## Fixed effects:
##                    Estimate Std. Error t value
```

```
## as.factor(Rater)1   1.6876     0.1212   13.92
## as.factor(Rater)2   2.1150     0.1237   17.09
## as.factor(Rater)3   1.8914     0.1225   15.45
##
## Correlation of Fixed Effects:
##             a.(R)1 a.(R)2
## as.fctr(R)2 0.249
## as.fctr(R)3 0.247  0.250
```

```
head(ranef(tmp), 10)
```

```
## $Artifact
##     (Intercept)
## 100  0.84082008
## 101 -0.44048526
## 102 -0.44048526
## 103  0.20016741
## 104 -0.44048526
## 105 -0.44048526
## 106  0.20016741
## 107 -0.44048526
## 111 -0.44048526
## 112 -0.44048526
## 113 -0.44048526
## 114 -0.44048526
## 115 -0.44048526
## 116 -0.44048526
## 117 -0.44048526
## 118 -0.44048526
## 13   0.06957550
## 15   0.71022817
## 16   0.71022817
## 17   0.06957550
## 21   0.71022817
## 22   0.71022817
## 23  -0.57107717
## 24   0.06957550
## 25   0.71022817
## 26  -0.57107717
## 27   0.06957550
## 28  -0.57107717
## 32   0.71022817
## 33   0.06957550
## 34   0.71022817
## 35  -0.57107717
## 36   0.06957550
## 37   0.71022817
## 38   0.06957550
## 39  -0.57107717
## 40   0.06957550
## 47   0.56696189
## 48   0.56696189
## 49  -0.71434345
## 53   1.20761456
## 54  -0.71434345
```

```
## 55  -0.07369078
## 56   0.56696189
## 57  -0.71434345
## 6   -0.57107717
## 61  -0.07369078
## 62   1.20761456
## 63   0.56696189
## 64   0.56696189
## 65   0.56696189
## 66   0.56696189
## 67  -0.71434345
## 68   0.56696189
## 7   -0.57107717
## 72  -0.07369078
## 73  -0.71434345
## 74  -0.71434345
## 75  -0.07369078
## 76  -0.07369078
## 77  -0.07369078
## 78   0.56696189
## 79  -0.07369078
## 8   -0.57107717
## 84   0.20016741
## 85   0.84082008
## 86   0.20016741
## 87   0.20016741
## 88   0.84082008
## 9   -0.57107717
## 92  -0.44048526
## 93  -0.44048526
## 94   0.84082008
## 95   0.20016741
## 96   0.20016741
## O1  -0.47571589
## O10 -0.47571589
## O11 -0.75654313
## O12 -0.19488866
## O13  0.08593858
## O2  -0.47571589
## O3   0.36676581
## O4   0.08593858
## O5   0.92842028
## O6  -0.75654313
## O7   0.36676581
## O8   0.08593858
## O9  -0.75654313
```

InterpRes

```
fla <- formula(model.formula.alldata[["InterpRes"]])
tmp <- lmer(fla,data=tall.nonmissing[tall.nonmissing$Rubric=="InterpRes",])
round(summary(tmp)$coef,2)
```

```
##                    Estimate Std. Error t value
## as.factor(Rater)1     2.70       0.09   30.34
```

```
## as.factor(Rater)2     2.59       0.09   29.01
## as.factor(Rater)3     2.14       0.09   23.70
```

```
tmp.single_intercept <- update(tmp, . ~ . + 1 - as.factor(Rater))
anova(tmp.single_intercept,tmp)
```

```
## refitting model(s) with ML (instead of REML)

## Data: tall.nonmissing[tall.nonmissing$Rubric == "InterpRes", ]
## Models:
## tmp.single_intercept: as.numeric(Rating) ~ (1 | Artifact)
## tmp: as.numeric(Rating) ~ as.factor(Rater) + (1 | Artifact) - 1
##                      npar    AIC    BIC   logLik deviance  Chisq Df Pr(>Chisq)
## tmp.single_intercept    3 218.53 226.79 -106.263   212.53
## tmp                     5 200.66 214.43  -95.331   190.66 21.864  2  1.787e-05
##
## tmp.single_intercept
## tmp                  ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
m0 <- tmp                                      ## Null hypothesis
mA <- update(m0, . ~ . + (as.factor(Rater)|Artifact))   ## Alternative hypotheses
```

```
## Error: number of observations (=116) <= number of random effects (=270) for term (as.factor(Rater) |
```

```
m  <- update(mA, . ~ . - (1|Artifact))          ## Model with only the new R.E.
```

```
## Error in h(simpleError(msg, call)): error in evaluating the argument 'object' in selecting a method f
```

```
exactRLRT(m0=m0,mA=mA,m=m)
```

```
## Error in exactRLRT(m0 = m0, mA = mA, m = m): object 'm' not found
```

```
summary(tmp)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: as.numeric(Rating) ~ as.factor(Rater) + (1 | Artifact) - 1
##    Data: tall.nonmissing[tall.nonmissing$Rubric == "InterpRes", ]
##
## REML criterion at convergence: 199.7
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.5317 -0.7627  0.2635  0.6614  2.6535
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Artifact (Intercept) 0.06224  0.2495
##  Residual             0.25250  0.5025
## Number of obs: 116, groups:  Artifact, 90
##
## Fixed effects:
##                    Estimate Std. Error t value
## as.factor(Rater)1  2.70421    0.08912   30.34
## as.factor(Rater)2  2.58574    0.08912   29.01
## as.factor(Rater)3  2.13918    0.09027   23.70
##
```

```
## Correlation of Fixed Effects:
##              a.(R)1 a.(R)2
## as.fctr(R)2 0.061
## as.fctr(R)3 0.062  0.062
```

```r
head(ranef(tmp), 10)
```

```
## $Artifact
##     (Intercept)
## 100  0.05848973
## 101 -0.13925357
## 102  0.05848973
## 103  0.05848973
## 104  0.05848973
## 105 -0.13925357
## 106 -0.13925357
## 107  0.05848973
## 111 -0.13925357
## 112  0.05848973
## 113  0.05848973
## 114  0.05848973
## 115  0.05848973
## 116 -0.13925357
## 117  0.05848973
## 118  0.05848973
## 13  -0.22526567
## 15  -0.02752238
## 16   0.17022092
## 17  -0.02752238
## 21   0.17022092
## 22   0.17022092
## 23  -0.22526567
## 24  -0.02752238
## 25  -0.22526567
## 26  -0.02752238
## 27  -0.02752238
## 28  -0.22526567
## 32   0.17022092
## 33  -0.02752238
## 34   0.17022092
## 35  -0.02752238
## 36  -0.02752238
## 37  -0.02752238
## 38  -0.02752238
## 39  -0.02752238
## 40  -0.02752238
## 45  -0.11582665
## 46  -0.11582665
## 47  -0.11582665
## 48   0.08191665
## 49   0.08191665
## 53   0.08191665
## 54  -0.11582665
## 55   0.08191665
## 56  -0.11582665
```

```
## 57  -0.11582665
## 6   -0.02752238
## 61  -0.11582665
## 62   0.08191665
## 63   0.08191665
## 64   0.08191665
## 65  -0.31356994
## 66   0.08191665
## 67   0.08191665
## 68   0.08191665
## 7   -0.02752238
## 72   0.08191665
## 73  -0.11582665
## 74   0.08191665
## 75   0.08191665
## 76  -0.11582665
## 77   0.08191665
## 78   0.08191665
## 79   0.08191665
## 8   -0.02752238
## 84   0.05848973
## 85   0.05848973
## 86   0.05848973
## 87  -0.13925357
## 88   0.05848973
## 9   -0.02752238
## 92   0.05848973
## 93   0.05848973
## 94   0.05848973
## 95   0.05848973
## 96   0.05848973
## O1   0.08089221
## O10  0.08089221
## O11  0.22259425
## O12 -0.06080983
## O13 -0.20251187
## O2  -0.06080983
## O3   0.08089221
## O4   0.22259425
## O5   0.08089221
## O6  -0.20251187
## O7   0.08089221
## O8  -0.34421391
## O9   0.22259425
```

VisOrg

```
fla <- formula(model.formula.alldata[["VisOrg"]])
tmp <- lmer(fla,data=tall.nonmissing[tall.nonmissing$Rubric=="VisOrg",])
round(summary(tmp)$coef,2)
```

```
##                    Estimate Std. Error t value
## as.factor(Rater)1      2.39        0.1   24.37
## as.factor(Rater)2      2.65        0.1   27.62
## as.factor(Rater)3      2.29        0.1   23.58
```

```r
tmp.single_intercept <- update(tmp, . ~ . + 1 - as.factor(Rater))
anova(tmp.single_intercept,tmp)
```

```
## refitting model(s) with ML (instead of REML)

## Data: tall.nonmissing[tall.nonmissing$Rubric == "VisOrg", ]
## Models:
## tmp.single_intercept: as.numeric(Rating) ~ (1 | Artifact)
## tmp: as.numeric(Rating) ~ as.factor(Rater) + (1 | Artifact) - 1
##                      npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## tmp.single_intercept    3 225.72 233.92 -109.86   219.72
## tmp                     5 219.49 233.17 -104.74   209.49 10.231  2   0.006002
##
## tmp.single_intercept
## tmp                  **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
m0 <- tmp                                        ## Null hypothesis
mA <- update(m0, . ~ . + (as.factor(Rater)|Artifact))   ## Alternative hypotheses
```

```
## Error: number of observations (=114) <= number of random effects (=264) for term (as.factor(Rater) |
```

```r
m  <- update(mA, . ~ . - (1|Artifact))          ## Model with only the new R.E.
```

```
## Error in h(simpleError(msg, call)): error in evaluating the argument 'object' in selecting a method f
```

```r
exactRLRT(m0=m0,mA=mA,m=m)
```

```
## Error in exactRLRT(m0 = m0, mA = mA, m = m): object 'm' not found
```

```r
summary(tmp)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: as.numeric(Rating) ~ as.factor(Rater) + (1 | Artifact) - 1
##    Data: tall.nonmissing[tall.nonmissing$Rubric == "VisOrg", ]
##
## REML criterion at convergence: 218.2
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.5099 -0.3379 -0.2492  0.3923  1.8404
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Artifact (Intercept) 0.2931   0.5414
##  Residual             0.1472   0.3837
## Number of obs: 114, groups:  Artifact, 88
##
## Fixed effects:
##                   Estimate Std. Error t value
## as.factor(Rater)1  2.38606    0.09791   24.37
## as.factor(Rater)2  2.65103    0.09599   27.62
## as.factor(Rater)3  2.28595    0.09694   23.58
##
## Correlation of Fixed Effects:
##             a.(R)1 a.(R)2
```

```
## as.fctr(R)2 0.266
## as.fctr(R)3 0.268  0.265
```

```r
head(ranef(tmp), 10)
```

```
## $Artifact
##     (Intercept)
## 101  0.40865880
## 103 -0.25697503
## 104 -0.25697503
## 105 -0.25697503
## 106 -0.25697503
## 107 -0.25697503
## 111 -0.25697503
## 112  0.40865880
## 113 -0.25697503
## 114 -0.25697503
## 115  0.40865880
## 116  0.40865880
## 117  1.07429264
## 118  0.40865880
## 13  -0.85596864
## 15   1.14093286
## 16   0.47529903
## 17  -0.19033481
## 21   0.47529903
## 22  -0.19033481
## 23  -0.85596864
## 24  -0.19033481
## 25  -0.19033481
## 26  -0.19033481
## 27  -0.19033481
## 28  -0.85596864
## 32   0.47529903
## 33  -0.19033481
## 34  -0.19033481
## 35   0.47529903
## 36  -0.19033481
## 37  -0.19033481
## 38   0.47529903
## 39  -0.19033481
## 40  -0.19033481
## 45  -0.43334477
## 46  -0.43334477
## 47  -1.09897860
## 48  -0.43334477
## 49   0.89792290
## 53   0.23228907
## 54  -0.43334477
## 55   0.23228907
## 56   0.23228907
## 57   0.23228907
## 6   -0.19033481
## 61   0.23228907
## 62   0.89792290
```

```
## 63   0.23228907
## 64   0.23228907
## 65   0.23228907
## 66   0.23228907
## 67   0.23228907
## 68   0.23228907
## 7   -0.19033481
## 72   0.23228907
## 73  -0.43334477
## 74   0.23228907
## 75  -0.43334477
## 76  -0.43334477
## 77   0.23228907
## 78   0.23228907
## 79   0.23228907
## 8   -0.19033481
## 84   0.40865880
## 85   0.40865880
## 86  -0.25697503
## 87  -0.25697503
## 88   1.07429264
## 9   -0.19033481
## 92  -0.25697503
## 93  -0.25697503
## 94   0.40865880
## 95  -0.25697503
## 96   0.40865880
## O1  -0.37775783
## O10 -0.09223341
## O11 -0.37775783
## O12 -0.37775783
## O13  0.19329100
## O2  -0.09223341
## O3   0.19329100
## O4  -0.09223341
## O5  -0.37775783
## O6  -0.09223341
## O7   0.47881542
## O8  -1.23433107
## O9   0.47881542
```

InitEDA

```
fla <- formula(model.formula.alldata[["InitEDA"]])
tmp <- lmer(fla,data=tall.nonmissing[tall.nonmissing$Rubric=="InitEDA",])
round(summary(tmp)$coef,2)
```

```
##             Estimate Std. Error t value
## (Intercept)     2.44       0.08    32.4
```

```
summary(tmp)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: as.numeric(Rating) ~ (1 | Artifact)
##    Data: tall.nonmissing[tall.nonmissing$Rubric == "InitEDA", ]
##
```

```
## REML criterion at convergence: 239
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.8889 -0.3391 -0.1427  0.4276  1.6035
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Artifact (Intercept) 0.3651   0.6042
##  Residual             0.1655   0.4068
## Number of obs: 116, groups:  Artifact, 90
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  2.44226    0.07537    32.4
```

```
head(ranef(tmp), 10)
```

```
## $Artifact
##     (Intercept)
## 100 -0.30430225
## 101  0.38376223
## 102 -0.30430225
## 103  0.38376223
## 104  0.38376223
## 105 -0.30430225
## 106 -0.99236674
## 107 -0.30430225
## 111 -0.30430225
## 112  0.38376223
## 113 -0.99236674
## 114  0.38376223
## 115  0.38376223
## 116 -0.30430225
## 117 -0.30430225
## 118 -0.30430225
## 13  -0.99236674
## 15   0.38376223
## 16   1.07182672
## 17  -0.30430225
## 21   1.07182672
## 22   0.38376223
## 23  -0.99236674
## 24  -0.30430225
## 25  -0.30430225
## 26  -0.30430225
## 27   0.38376223
## 28  -0.99236674
## 32   0.38376223
## 33   0.38376223
## 34  -0.30430225
## 35  -0.30430225
## 36  -0.30430225
## 37  -0.30430225
## 38  -0.30430225
```

```
## 39   0.38376223
## 40   0.38376223
## 45  -0.30430225
## 46   0.38376223
## 47  -0.30430225
## 48   1.07182672
## 49   0.38376223
## 53   0.38376223
## 54   0.38376223
## 55  -0.30430225
## 56  -0.30430225
## 57  -0.30430225
## 6   -0.30430225
## 61   0.38376223
## 62   1.07182672
## 63   0.38376223
## 64  -0.30430225
## 65  -0.30430225
## 66   1.07182672
## 67   0.38376223
## 68  -0.30430225
## 7    0.38376223
## 72   0.38376223
## 73  -0.99236674
## 74   0.38376223
## 75   0.38376223
## 76  -0.30430225
## 77  -0.30430225
## 78   0.38376223
## 79   0.38376223
## 8   -0.30430225
## 84  -0.30430225
## 85   0.38376223
## 86  -0.30430225
## 87  -0.99236674
## 88   0.38376223
## 9   -0.30430225
## 92  -0.30430225
## 93  -0.30430225
## 94   1.07182672
## 95   0.38376223
## 96   0.38376223
## O1   0.48452197
## O10 -0.09462545
## O11 -0.38419916
## O12 -0.38419916
## O13  0.19494826
## O2  -0.09462545
## O3  -0.09462545
## O4   0.19494826
## O5  -0.38419916
## O6  -0.67377287
## O7   0.48452197
## O8  -0.38419916
```

```
## 09   0.48452197
```

RsrchQ

```
fla <- formula(model.formula.alldata[["RsrchQ"]])
tmp <- lmer(fla,data=tall.nonmissing[tall.nonmissing$Rubric=="RsrchQ",])
round(summary(tmp)$coef,2)
```

```
##             Estimate Std. Error t value
## (Intercept)     2.35       0.06   40.59
```

```
summary(tmp)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: as.numeric(Rating) ~ (1 | Artifact)
##    Data: tall.nonmissing[tall.nonmissing$Rubric == "RsrchQ", ]
##
## REML criterion at convergence: 209.1
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.2694 -0.5285 -0.3736  0.9743  2.4770
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Artifact (Intercept) 0.07276  0.2697
##  Residual             0.27825  0.5275
## Number of obs: 116, groups:  Artifact, 90
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  2.35169    0.05794   40.59
```

```
head(ranef(tmp), 10)
```

```
## $Artifact
##     (Intercept)
## 100 -0.072903664
## 101 -0.280199221
## 102 -0.280199221
## 103 -0.072903664
## 104 -0.072903664
## 105 -0.072903664
## 106  0.134391893
## 107  0.134391893
## 111 -0.072903664
## 112  0.134391893
## 113 -0.072903664
## 114 -0.072903664
## 115  0.134391893
## 116 -0.072903664
## 117 -0.072903664
## 118 -0.072903664
## 13  -0.072903664
## 15  -0.072903664
## 16  -0.072903664
## 17   0.134391893
```

```
## 21   0.134391893
## 22   0.134391893
## 23  -0.072903664
## 24  -0.072903664
## 25  -0.072903664
## 26  -0.072903664
## 27  -0.072903664
## 28  -0.280199221
## 32   0.134391893
## 33  -0.072903664
## 34  -0.072903664
## 35  -0.072903664
## 36  -0.072903664
## 37  -0.072903664
## 38  -0.072903664
## 39   0.134391893
## 40  -0.072903664
## 45  -0.072903664
## 46  -0.072903664
## 47   0.134391893
## 48   0.134391893
## 49   0.134391893
## 53   0.134391893
## 54  -0.280199221
## 55   0.134391893
## 56  -0.072903664
## 57  -0.072903664
## 6   -0.072903664
## 61  -0.072903664
## 62   0.134391893
## 63   0.134391893
## 64  -0.072903664
## 65   0.134391893
## 66   0.134391893
## 67   0.134391893
## 68   0.134391893
## 7   -0.072903664
## 72  -0.072903664
## 73  -0.072903664
## 74  -0.072903664
## 75  -0.072903664
## 76  -0.072903664
## 77  -0.072903664
## 78   0.134391893
## 79   0.134391893
## 8   -0.072903664
## 84   0.134391893
## 85   0.341687450
## 86   0.134391893
## 87   0.134391893
## 88   0.134391893
## 9    0.134391893
## 92   0.134391893
## 93   0.134391893
```

```
## 94   0.134391893
## 95  -0.072903664
## 96   0.134391893
## O1  -0.008069777
## O10 -0.008069777
## O11  0.285012168
## O12 -0.154610749
## O13 -0.154610749
## O2  -0.154610749
## O3   0.138471195
## O4  -0.008069777
## O5   0.138471195
## O6  -0.154610749
## O7   0.138471195
## O8  -0.301151721
## O9  -0.154610749
```

TxtOrg

```
fla <- formula(model.formula.alldata[["TxtOrg"]])
tmp <- lmer(fla,data=tall.nonmissing[tall.nonmissing$Rubric=="TxtOrg",])
round(summary(tmp)$coef,2)
```

```
##             Estimate Std. Error t value
## (Intercept)     2.59       0.07   37.93
```

```
summary(tmp)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: as.numeric(Rating) ~ (1 | Artifact)
##    Data: tall.nonmissing[tall.nonmissing$Rubric == "TxtOrg", ]
##
## REML criterion at convergence: 247.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.3557 -0.7550  0.3834  0.5302  2.4132
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Artifact (Intercept) 0.09371  0.3061
##  Residual             0.39573  0.6291
## Number of obs: 116, groups:  Artifact, 90
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  2.58745    0.06821   37.93
```

Above are the final models for rubric specific models.

Trying to add fixed effects, interactions, and new random effects to the "combined" model Rating ~ 1 + (0 + Rubric|Artifact), using all the data.

```
comb.0 <- lmer(as.numeric(Rating) ~ 1 + (0 + Rubric | Artifact),
               data=tall.nonmissing)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## unable to evaluate scaled gradient

## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge: degenerate Hessian with 1 negative eigenvalues
```

```
summary(comb.0)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: as.numeric(Rating) ~ 1 + (0 + Rubric | Artifact)
##    Data: tall.nonmissing
##
## REML criterion at convergence: 1469.3
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.0208 -0.4915 -0.0836  0.5305  3.7746
##
## Random effects:
##  Groups   Name            Variance Std.Dev. Corr
##  Artifact RubricCritDes   0.64732  0.8046
##           RubricInitEDA   0.38336  0.6192   0.26
##           RubricInterpRes 0.25633  0.5063   0.00 0.79
##           RubricRsrchQ    0.17393  0.4170   0.38 0.50 0.74
##           RubricSelMeth   0.09595  0.3098   0.56 0.37 0.41 0.26
##           RubricTxtOrg    0.40442  0.6359   0.03 0.69 0.80 0.64 0.24
##           RubricVisOrg    0.32324  0.5685   0.17 0.78 0.76 0.60 0.29 0.80
##  Residual                 0.19490  0.4415
## Number of obs: 808, groups:  Artifact, 90
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  2.23179    0.04014   55.59
## optimizer (nloptwrap) convergence code: 0 (OK)
## unable to evaluate scaled gradient
## Model failed to converge: degenerate  Hessian with 1 negative eigenvalues
```

```
comb.full <- update(comb.0, . ~ . + as.factor(Rater) + Semester +
                      Sex + Repeated + Rubric)
```

```
## boundary (singular) fit: see ?isSingular
```

```
summary(comb.full)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: as.numeric(Rating) ~ (0 + Rubric | Artifact) + as.factor(Rater) +
##     Semester + Sex + Repeated + Rubric
##    Data: tall.nonmissing
##
## REML criterion at convergence: 1427.8
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.1097 -0.5106 -0.0242  0.5287  3.7537
##
## Random effects:
##  Groups   Name            Variance Std.Dev. Corr
```

```
##  Artifact RubricCritDes   0.55757  0.7467
##           RubricInitEDA   0.35009  0.5917   0.46
##           RubricInterpRes 0.17260  0.4155   0.23 0.76
##           RubricRsrchQ    0.16821  0.4101   0.59 0.44 0.72
##           RubricSelMeth   0.06709  0.2590   0.39 0.61 0.75 0.42
##           RubricTxtOrg    0.25908  0.5090   0.34 0.62 0.74 0.55 0.67
##           RubricVisOrg    0.26135  0.5112   0.35 0.74 0.67 0.53 0.41 0.78
##  Residual                 0.19129  0.4374
## Number of obs: 808, groups:  Artifact, 90
##
## Fixed effects:
##                    Estimate Std. Error t value
## (Intercept)        2.011815   0.109810  18.321
## as.factor(Rater)2  0.001785   0.055145   0.032
## as.factor(Rater)3 -0.175088   0.055302  -3.166
## SemesterS19       -0.174173   0.088015  -1.979
## SexM               0.010636   0.081427   0.131
## Repeated          -0.070713   0.098669  -0.717
## RubricInitEDA      0.548777   0.096480   5.688
## RubricInterpRes    0.587266   0.101491   5.786
## RubricRsrchQ       0.462880   0.088265   5.244
## RubricSelMeth      0.166301   0.095016   1.750
## RubricTxtOrg       0.694646   0.099823   6.959
## RubricVisOrg       0.528827   0.100012   5.288
##
## Correlation of Fixed Effects:
##             (Intr) a.(R)2 a.(R)3 SmsS19 SexM   Repetd RbIEDA RbrcIR RbrcRQ
## as.fctr(R)2 -0.244
## as.fctr(R)3 -0.237  0.499
## SemesterS19 -0.358  0.008  0.000
## SexM        -0.396 -0.026 -0.035  0.301
## Repeated    -0.153  0.001 -0.003  0.079  0.009
## RubrcIntEDA -0.558 -0.001  0.000 -0.002  0.001  0.008
## RbrcIntrpRs -0.664 -0.001  0.000 -0.002  0.001 -0.010  0.739
## RubrcRsrchQ -0.631 -0.001  0.000 -0.002  0.001 -0.038  0.592  0.761
## RubricSlMth -0.692 -0.001  0.000 -0.002  0.001 -0.087  0.665  0.779  0.695
## RubrcTxtOrg -0.615 -0.001  0.000 -0.002  0.001  0.004  0.676  0.762  0.682
## RubricVsOrg -0.608 -0.003 -0.002 -0.004  0.001 -0.021  0.719  0.743  0.674
##             RbrcSM RbrcTO
## as.fctr(R)2
## as.fctr(R)3
## SemesterS19
## SexM
## Repeated
## RubrcIntEDA
## RbrcIntrpRs
## RubrcRsrchQ
## RubricSlMth
## RubrcTxtOrg  0.727
## RubricVsOrg  0.681  0.757
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see ?isSingular
```

some of the random effects are highly correlated with one another. We can see this in the “Random effects". The random effects for VisOrg and TxtOrg seem highly correlated with each other and with everything except for the random effect for SelMeth. The random effects for InterpRes and InitEDA are highly correlated. The random effects for RsrchQ and InterpRes are highly correlated etc. In some ways we should not be surprised: these rubrics all represent features of a good research report, and we would expect that if someoneis good at one or two of these features, they are probably good at the others.

```
comb.back_elim <- fitLMER.fnc(comb.full, log.file.name = FALSE)
```

```
## Warning in fitLMER.fnc(comb.full, log.file.name = FALSE): Argument "ran.effects" is empty, which mea
## TRUE

## ======================================================
## ===              backfitting fixed effects         ===
## ======================================================
## processing model terms of interaction level 1
##   iteration 1
##     p-value for term "Sex" = 0.8886 >= 0.05
##     not part of higher-order interaction
##     removing term
##   iteration 2
##     p-value for term "Repeated" = 0.0934 >= 0.05
##     not part of higher-order interaction

## boundary (singular) fit: see ?isSingular

##     removing term
## pruning random effects structure ...
##   nothing to prune
## ======================================================
## ===            forwardfitting random effects       ===
## ======================================================
##  ===       random slopes       ===
## ======================================================
## ===            re-backfitting fixed effects        ===
## ======================================================
## processing model terms of interaction level 1
##   all terms of interaction level 1 significant
## resetting REML to TRUE

## boundary (singular) fit: see ?isSingular

## pruning random effects structure ...
##   nothing to prune
```

```
summary(comb.back_elim)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: as.numeric(Rating) ~ (0 + Rubric | Artifact) + as.factor(Rater) +
##     Semester + Rubric
##    Data: tall.nonmissing
##
## REML criterion at convergence: 1422.3
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.1240 -0.5150 -0.0166  0.5276  3.7365
##
```

```
## Random effects:
##  Groups   Name             Variance Std.Dev. Corr
##  Artifact RubricCritDes    0.55945  0.7480
##           RubricInitEDA    0.34861  0.5904   0.46
##           RubricInterpRes  0.16683  0.4084   0.23 0.76
##           RubricRsrchQ     0.16617  0.4076   0.59 0.44 0.71
##           RubricSelMeth    0.06409  0.2532   0.40 0.61 0.74 0.41
##           RubricTxtOrg     0.25251  0.5025   0.34 0.61 0.72 0.54 0.65
##           RubricVisOrg     0.26240  0.5122   0.35 0.74 0.67 0.53 0.41 0.77
##  Residual                  0.19124  0.4373
## Number of obs: 808, groups:  Artifact, 90
##
## Fixed effects:
##                      Estimate Std. Error t value
## (Intercept)         2.0069559  0.0994799  20.174
## as.factor(Rater)2   0.0001728  0.0549737   0.003
## as.factor(Rater)3  -0.1771897  0.0551181  -3.215
## SemesterS19        -0.1724092  0.0828434  -2.081
## RubricInitEDA       0.5491728  0.0964883   5.692
## RubricInterpRes     0.5866116  0.1014621   5.782
## RubricRsrchQ        0.4605262  0.0881655   5.223
## RubricSelMeth       0.1607441  0.0945337   1.700
## RubricTxtOrg        0.6946497  0.0998581   6.956
## RubricVisOrg        0.5276703  0.0999783   5.278
##
## Correlation of Fixed Effects:
##             (Intr) a.(R)2 a.(R)3 SmsS19 RbIEDA RbrcIR RbrcRQ RbrcSM RbrcTO
## as.fctr(R)2 -0.279
## as.fctr(R)3 -0.276  0.499
## SemesterS19 -0.262  0.017  0.012
## RubrcIntEDA -0.616 -0.001  0.000 -0.003
## RbrcIntrpRs -0.739 -0.001  0.000 -0.001  0.739
## RubrcRsrchQ -0.705 -0.001  0.000  0.001  0.593  0.761
## RubricSlMth -0.784 -0.001  0.000  0.005  0.668  0.782  0.694
## RubrcTxtOrg -0.683 -0.001  0.000 -0.003  0.677  0.761  0.682  0.730
## RubricVsOrg -0.675 -0.003 -0.002 -0.002  0.719  0.743  0.673  0.681  0.756
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see ?isSingular
```

```r
comb.inter <- update(comb.back_elim, . ~ . + as.factor(Rater)*Semester*Rubric)
```

```
## boundary (singular) fit: see ?isSingular
```

```r
ss <- getME(comb.inter,c("theta","fixef"))
comb.inter.u<- update(comb.inter,start=ss,
             control=lmerControl(optimizer="bobyqa",
                                 optCtrl=list(maxfun=2e5)))
```

```
## boundary (singular) fit: see ?isSingular
```

```r
summary(comb.inter.u)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: as.numeric(Rating) ~ (0 + Rubric | Artifact) + as.factor(Rater) +
##     Semester + Rubric + as.factor(Rater):Semester + as.factor(Rater):Rubric +
##     Semester:Rubric + as.factor(Rater):Semester:Rubric
```

```
##    Data: tall.nonmissing
## Control: lmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 2e+05))
##
## REML criterion at convergence: 1422.1
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.9245 -0.5150 -0.0667  0.5128  3.6299
##
## Random effects:
##  Groups   Name            Variance Std.Dev. Corr
##  Artifact RubricCritDes   0.48896  0.6993
##           RubricInitEDA   0.35171  0.5930   0.41
##           RubricInterpRes 0.14396  0.3794   0.32 0.81
##           RubricRsrchQ    0.16398  0.4049   0.65 0.43 0.73
##           RubricSelMeth   0.06238  0.2498   0.45 0.65 0.79 0.49
##           RubricTxtOrg    0.25078  0.5008   0.44 0.65 0.71 0.59 0.61
##           RubricVisOrg    0.26281  0.5126   0.35 0.72 0.68 0.58 0.33 0.79
##  Residual                 0.18985  0.4357
## Number of obs: 808, groups:  Artifact, 90
##
## Fixed effects:
##                                             Estimate Std. Error t value
## (Intercept)                                  1.73980    0.13709  12.691
## as.factor(Rater)2                            0.30221    0.15574   1.941
## as.factor(Rater)3                            0.23697    0.15650   1.514
## SemesterS19                                 -0.12772    0.25161  -0.508
## RubricInitEDA                                0.76684    0.16586   4.623
## RubricInterpRes                              0.97905    0.16269   6.018
## RubricRsrchQ                                 0.71182    0.14821   4.803
## RubricSelMeth                                0.46289    0.15579   2.971
## RubricTxtOrg                                 1.01067    0.16085   6.283
## RubricVisOrg                                 0.63047    0.16851   3.741
## as.factor(Rater)2:SemesterS19                0.27804    0.30893   0.900
## as.factor(Rater)3:SemesterS19               -0.07187    0.30225  -0.238
## as.factor(Rater)2:RubricInitEDA             -0.32639    0.20493  -1.593
## as.factor(Rater)3:RubricInitEDA             -0.37587    0.20618  -1.823
## as.factor(Rater)2:RubricInterpRes           -0.47112    0.20171  -2.336
## as.factor(Rater)3:RubricInterpRes           -0.71186    0.20297  -3.507
## as.factor(Rater)2:RubricRsrchQ              -0.44705    0.19025  -2.350
## as.factor(Rater)3:RubricRsrchQ              -0.47500    0.19160  -2.479
## as.factor(Rater)2:RubricSelMeth             -0.30083    0.19434  -1.548
## as.factor(Rater)3:RubricSelMeth             -0.36572    0.19563  -1.869
## as.factor(Rater)2:RubricTxtOrg              -0.44634    0.20101  -2.220
## as.factor(Rater)3:RubricTxtOrg              -0.40651    0.20230  -2.009
## as.factor(Rater)2:RubricVisOrg               0.02255    0.20665   0.109
## as.factor(Rater)3:RubricVisOrg              -0.27119    0.20788  -1.305
## SemesterS19:RubricInitEDA                   -0.05173    0.30280  -0.171
## SemesterS19:RubricInterpRes                  0.12378    0.29696   0.417
## SemesterS19:RubricRsrchQ                     0.13283    0.26943   0.493
## SemesterS19:RubricSelMeth                   -0.08971    0.28402  -0.316
## SemesterS19:RubricTxtOrg                     0.16478    0.29329   0.562
## SemesterS19:RubricVisOrg                     0.16119    0.30471   0.529
## as.factor(Rater)2:SemesterS19:RubricInitEDA  0.01089    0.39754   0.027
```

```
## as.factor(Rater)3:SemesterS19:RubricInitEDA    0.25207    0.39151   0.644
## as.factor(Rater)2:SemesterS19:RubricInterpRes -0.27037    0.38982  -0.694
## as.factor(Rater)3:SemesterS19:RubricInterpRes -0.15207    0.38461  -0.395
## as.factor(Rater)2:SemesterS19:RubricRsrchQ    -0.22860    0.36659  -0.624
## as.factor(Rater)3:SemesterS19:RubricRsrchQ     0.35485    0.35919   0.988
## as.factor(Rater)2:SemesterS19:RubricSelMeth   -0.41331    0.37498  -1.102
## as.factor(Rater)3:SemesterS19:RubricSelMeth   -0.19456    0.36912  -0.527
## as.factor(Rater)2:SemesterS19:RubricTxtOrg    -0.55532    0.38875  -1.429
## as.factor(Rater)3:SemesterS19:RubricTxtOrg    -0.31424    0.38267  -0.821
## as.factor(Rater)2:SemesterS19:RubricVisOrg    -0.62601    0.39828  -1.572
## as.factor(Rater)3:SemesterS19:RubricVisOrg    -0.20005    0.39282  -0.509

##
## Correlation matrix not shown by default, as p = 42 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it

## optimizer (bobyqa) convergence code: 0 (OK)
## boundary (singular) fit: see ?isSingular
```

```
comb.inter_elim <- fitLMER.fnc(comb.inter.u, log.file.name = FALSE)
```

```
## Warning in fitLMER.fnc(comb.inter.u, log.file.name = FALSE): Argument "ran.effects" is empty, which r
## TRUE

## ========================================================
## ===              backfitting fixed effects         ===
## ========================================================
## processing model terms of interaction level 3
##   iteration 1
##     p-value for term "as.factor(Rater):Semester:Rubric" = 0.5468 >= 0.05
##     not part of higher-order interaction

## boundary (singular) fit: see ?isSingular

##     removing term
## processing model terms of interaction level 2
##   iteration 2
##     p-value for term "as.factor(Rater):Semester" = 0.5976 >= 0.05
##     not part of higher-order interaction

## boundary (singular) fit: see ?isSingular

##     removing term
##   iteration 3
##     p-value for term "Semester:Rubric" = 0.0773 >= 0.05
##     not part of higher-order interaction

## boundary (singular) fit: see ?isSingular

##     removing term
## processing model terms of interaction level 1
##   all terms of interaction level 1 significant
## pruning random effects structure ...
##   nothing to prune
## ========================================================
## ===            forwardfitting random effects       ===
## ========================================================
##  ===       random slopes       ===
```

```
## ========================================================
## ===           re-backfitting fixed effects          ===
## ========================================================
## processing model terms of interaction level 2
##   all terms of interaction level 2 significant
## processing model terms of interaction level 1
##   all terms of interaction level 1 significant
## resetting REML to TRUE
```

```
## boundary (singular) fit: see ?isSingular
```

```
## pruning random effects structure ...
##   nothing to prune
```

```
summary(comb.inter_elim)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: as.numeric(Rating) ~ (0 + Rubric | Artifact) + as.factor(Rater) +
##     Semester + Rubric + as.factor(Rater):Rubric
##    Data: tall.nonmissing
## Control: lmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 2e+05))
##
## REML criterion at convergence: 1417.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.9705 -0.5193 -0.0406  0.4920  3.5506
##
## Random effects:
##  Groups   Name            Variance Std.Dev. Corr
##  Artifact RubricCritDes   0.50798  0.7127
##           RubricInitEDA   0.35349  0.5946   0.44
##           RubricInterpRes 0.14982  0.3871   0.36 0.82
##           RubricRsrchQ    0.17843  0.4224   0.63 0.44 0.72
##           RubricSelMeth   0.06658  0.2580   0.43 0.61 0.74 0.37
##           RubricTxtOrg    0.25700  0.5070   0.42 0.63 0.72 0.54 0.63
##           RubricVisOrg    0.26213  0.5120   0.34 0.71 0.68 0.53 0.36 0.80
##  Residual                 0.18676  0.4322
## Number of obs: 808, groups:  Artifact, 90
##
## Fixed effects:
##                                  Estimate Std. Error t value
## (Intercept)                       1.75875    0.11843  14.850
## as.factor(Rater)2                 0.36487    0.13404   2.722
## as.factor(Rater)3                 0.21412    0.13358   1.603
## SemesterS19                      -0.17687    0.08247  -2.145
## RubricInitEDA                     0.74803    0.13743   5.443
## RubricInterpRes                   1.01438    0.13525   7.500
## RubricRsrchQ                      0.75095    0.12484   6.015
## RubricSelMeth                     0.42808    0.13100   3.268
## RubricTxtOrg                      1.04958    0.13560   7.740
## RubricVisOrg                      0.67274    0.14082   4.777
## as.factor(Rater)2:RubricInitEDA  -0.30903    0.17376  -1.779
## as.factor(Rater)3:RubricInitEDA  -0.29712    0.17362  -1.711
## as.factor(Rater)2:RubricInterpRes -0.53711    0.17100  -3.141
## as.factor(Rater)3:RubricInterpRes -0.75335    0.17101  -4.405
```

```
## as.factor(Rater)2:RubricRsrchQ    -0.50134    0.16280  -3.079
## as.factor(Rater)3:RubricRsrchQ    -0.37154    0.16252  -2.286
## as.factor(Rater)2:RubricSelMeth   -0.39637    0.16578  -2.391
## as.factor(Rater)3:RubricSelMeth   -0.41463    0.16571  -2.502
## as.factor(Rater)2:RubricTxtOrg    -0.58169    0.17200  -3.382
## as.factor(Rater)3:RubricTxtOrg    -0.48503    0.17191  -2.821
## as.factor(Rater)2:RubricVisOrg    -0.13559    0.17600  -0.770
## as.factor(Rater)3:RubricVisOrg    -0.32306    0.17599  -1.836
```

```
##
## Correlation matrix not shown by default, as p = 22 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it
```

```
## optimizer (bobyqa) convergence code: 0 (OK)
## boundary (singular) fit: see ?isSingular
```

```
formula(comb.inter.u)
```

```
## as.numeric(Rating) ~ (0 + Rubric | Artifact) + as.factor(Rater) +
##     Semester + Rubric + as.factor(Rater):Semester + as.factor(Rater):Rubric +
##     Semester:Rubric + as.factor(Rater):Semester:Rubric
```

```
formula(comb.inter_elim)
```

```
## as.numeric(Rating) ~ (0 + Rubric | Artifact) + as.factor(Rater) +
##     Semester + Rubric + as.factor(Rater):Rubric
```

```
formula(comb.back_elim)
```

```
## as.numeric(Rating) ~ (0 + Rubric | Artifact) + as.factor(Rater) +
##     Semester + Rubric
```

```
summary(comb.inter.u)$varcor
```

```
##  Groups   Name            Std.Dev. Corr
##  Artifact RubricCritDes   0.69926
##           RubricInitEDA   0.59305  0.414
##           RubricInterpRes 0.37942  0.322 0.806
##           RubricRsrchQ    0.40494  0.650 0.433 0.730
##           RubricSelMeth   0.24975  0.448 0.645 0.794 0.493
##           RubricTxtOrg    0.50078  0.442 0.646 0.711 0.591 0.610
##           RubricVisOrg    0.51265  0.351 0.723 0.679 0.583 0.333 0.788
##  Residual                 0.43572
```

```
summary(comb.inter_elim)$varcor
```

```
##  Groups   Name            Std.Dev. Corr
##  Artifact RubricCritDes   0.71272
##           RubricInitEDA   0.59455  0.442
##           RubricInterpRes 0.38707  0.356 0.822
##           RubricRsrchQ    0.42241  0.629 0.443 0.723
##           RubricSelMeth   0.25804  0.425 0.608 0.745 0.370
##           RubricTxtOrg    0.50695  0.422 0.634 0.715 0.536 0.627
##           RubricVisOrg    0.51199  0.342 0.712 0.680 0.528 0.365 0.802
##  Residual                 0.43216
```

```
summary(comb.back_elim)$varcor
```

```
##  Groups   Name            Std.Dev. Corr
##  Artifact RubricCritDes   0.74796
##           RubricInitEDA   0.59043  0.464
##           RubricInterpRes 0.40844  0.232 0.757
##           RubricRsrchQ    0.40764  0.589 0.438 0.713
##           RubricSelMeth   0.25317  0.401 0.610 0.742 0.407
##           RubricTxtOrg    0.50250  0.341 0.613 0.725 0.542 0.652
##           RubricVisOrg    0.51225  0.353 0.735 0.673 0.531 0.413 0.773
##  Residual                 0.43731
```

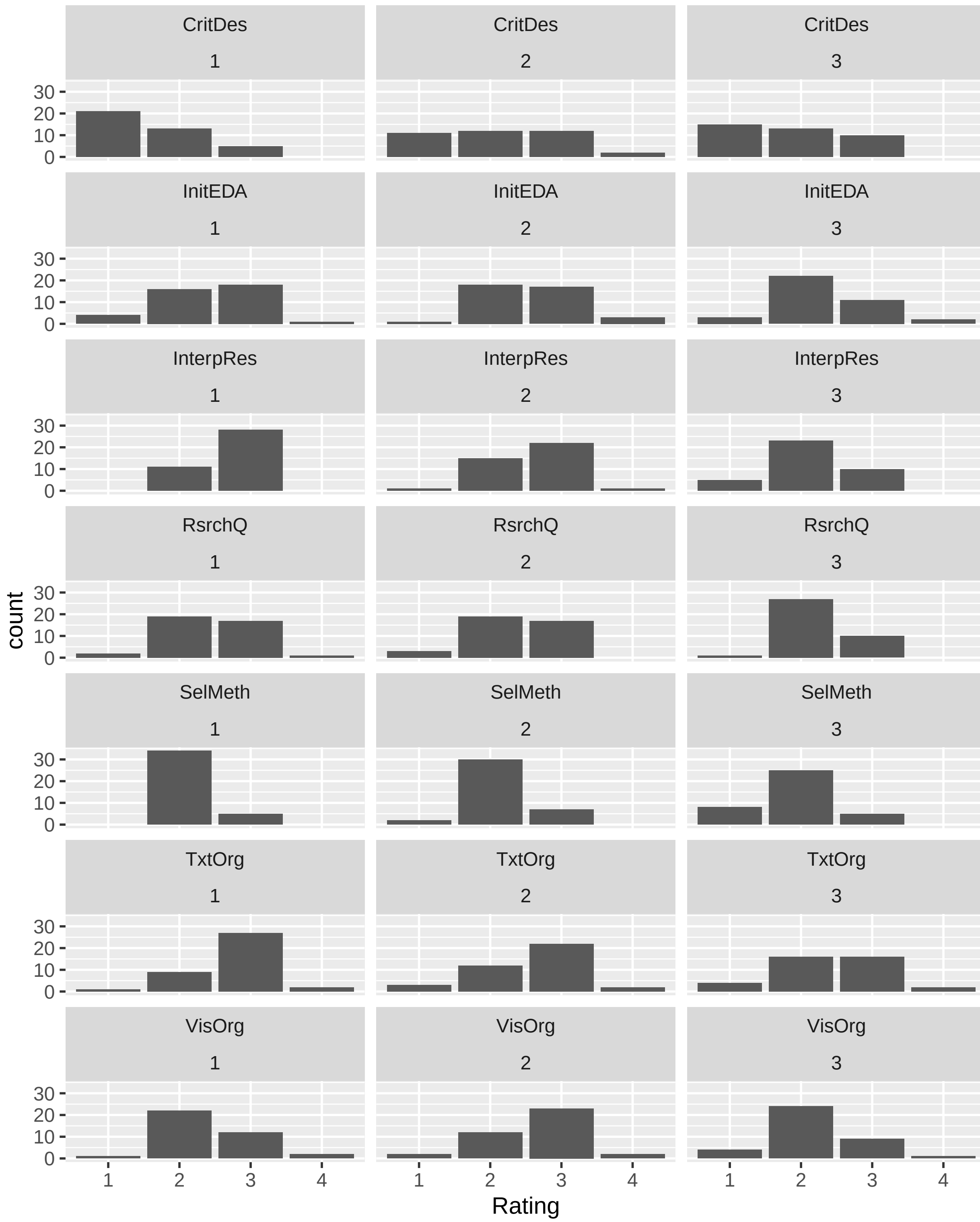
```
anova(comb.back_elim,comb.inter_elim,comb.inter.u)
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: tall.nonmissing
## Models:
## comb.back_elim: as.numeric(Rating) ~ (0 + Rubric | Artifact) + as.factor(Rater) + Semester + Rubric
## comb.inter_elim: as.numeric(Rating) ~ (0 + Rubric | Artifact) + as.factor(Rater) + Semester + Rubric
## comb.inter.u: as.numeric(Rating) ~ (0 + Rubric | Artifact) + as.factor(Rater) + Semester + Rubric + a
##                 npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## comb.back_elim    39 1462.2 1645.3 -692.11   1384.2
## comb.inter_elim   51 1452.6 1692.0 -675.28   1350.6 33.661 12  0.0007628 ***
## comb.inter.u      71 1469.4 1802.7 -663.68   1327.4 23.199 20  0.2791228
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

> the models are nested so we can use AIC, BIC or likelihod ratio (deviance) tests… AIC and the LRT agree on comb.inter_elim; BIC likes the simpler comb.back_elim.Interestingly, comb.inter_elim adds a rater x rubric interaction to the main-effects model comb.back_elim. This suggests that the raters do not all use the rubrics in the same way.

```
g <- ggplot(tall.nonmissing, aes(x=Rating)) +
  geom_bar() +
  facet_wrap( ~ Rubric + Rater, nrow=7)

g
```



it does look as if the 3 raters have different ways of scoring the 7 rubrics, so the interaction we found in comb.inter_elim makes sense. (Clearly it is not the case that one rater is simply more harsh than another, or something like that.

Adding some random effects

```r
m0 <- comb.inter_elim
mA <- lmer(as.numeric(Rating) ~ (0 + Rubric | Artifact) +
             (0 + as.factor(Rater) | Artifact) + as.factor(Rater) +
             Semester + Rubric + as.factor(Rater):Rubric, data=tall.nonmissing)
```

```
## boundary (singular) fit: see ?isSingular
```

```r
anova(m0,mA)
```

```
## refitting model(s) with ML (instead of REML)

## Warning in commonArgs(par, fn, control, environment()): maxfun < 10 *
## length(par)^2 is not recommended.

## Data: tall.nonmissing
## Models:
## m0: as.numeric(Rating) ~ (0 + Rubric | Artifact) + as.factor(Rater) + Semester + Rubric + as.factor(R
## mA: as.numeric(Rating) ~ (0 + Rubric | Artifact) + (0 + as.factor(Rater) | Artifact) + as.factor(Rate
##    npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## m0   51 1452.6 1692.0 -675.28   1350.6
## mA   57 1413.9 1681.5 -649.96   1299.9 50.644  6  3.492e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
## AIC and BIC both like including (0 + as.factor(Rater) | Artifact) in the model
```

```r
m0 <- comb.inter_elim
mA <- lmer(as.numeric(Rating) ~ (0 + Rubric | Artifact) +
             (0 + Semester | Artifact) + as.factor(Rater) +
             Semester + Rubric + as.factor(Rater):Rubric, data=tall.nonmissing)
```

```
## boundary (singular) fit: see ?isSingular
```

```r
anova(m0,mA)
```

```
## refitting model(s) with ML (instead of REML)

## Data: tall.nonmissing
## Models:
## m0: as.numeric(Rating) ~ (0 + Rubric | Artifact) + as.factor(Rater) + Semester + Rubric + as.factor(R
## mA: as.numeric(Rating) ~ (0 + Rubric | Artifact) + (0 + Semester | Artifact) + as.factor(Rater) + Sem
##    npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## m0   51 1452.6 1692.0 -675.28   1350.6
## mA   54 1456.5 1710.1 -674.28   1348.5 2.0025  3     0.5719
```

```r
##
## AIC and BIC do not like (0 + Semester | Artifact) in the model...
```

```r
m0 <- comb.inter_elim
mA <- lmer(as.numeric(Rating) ~ (0 + Rubric | Artifact) +
             (0 + as.factor(Rater) | Artifact) +
             (0 + as.factor(Rater):Rubric | Artifact) + as.factor(Rater) +
             Semester + Rubric + as.factor(Rater):Rubric, data=tall.nonmissing)
```

```
## Error: number of observations (=808) <= number of random effects (=1890) for term (0 + as.factor(Rate
```

```r
## anova(m0,mA)     -- Not needed!
##
## There are not enough observations to fit mA here, so we need not do any
## formal model comparison...
```

```r
## So, to summarize, the "final" model appears to be

comb.final <- lmer(as.numeric(Rating) ~ (0 + Rubric | Artifact) +
             (0 + as.factor(Rater) | Artifact) + as.factor(Rater) +
             Semester + Rubric + as.factor(Rater):Rubric, data=tall.nonmissing)
```

```
## boundary (singular) fit: see ?isSingular
```

```r
formula(comb.final)
```

```
## as.numeric(Rating) ~ (0 + Rubric | Artifact) + (0 + as.factor(Rater) |
##     Artifact) + as.factor(Rater) + Semester + Rubric + as.factor(Rater):Rubric
```

```r
summary(comb.final)$varcor
```

```
##  Groups     Name               Std.Dev. Corr
##  Artifact   RubricCritDes      0.70840
##             RubricInitEDA      0.56355   0.317
##             RubricInterpRes    0.31918   0.141  0.675
##             RubricRsrchQ       0.42267   0.498  0.194  0.540
##             RubricSelMeth      0.19489   0.145  0.227  0.378 -0.243
##             RubricTxtOrg       0.50038   0.267  0.438  0.367  0.306  0.216
##             RubricVisOrg       0.48950   0.177  0.503  0.442  0.294 -0.157
##  Artifact.1 as.factor(Rater)1 0.11376
##             as.factor(Rater)2 0.33400  -0.488
##             as.factor(Rater)3 0.30632   0.332  0.662
##  Residual                      0.36739
##
##
##
##
##
##
##
##   0.554
##
##
##
##
```

```r
summary(comb.final)$coef
```

```
##                                     Estimate Std. Error    t value
## (Intercept)                       1.75781939 0.11454398 15.3462402
## as.factor(Rater)2                 0.36632775 0.13983142  2.6197813
## as.factor(Rater)3                 0.19511357 0.12991876  1.5018122
## SemesterS19                      -0.15847010 0.07668583 -2.0664846
## RubricInitEDA                     0.73917906 0.13036874  5.6699102
## RubricInterpRes                   0.99120641 0.12818485  7.7326328
## RubricRsrchQ                      0.72530401 0.11838651  6.1265764
## RubricSelMeth                     0.40998310 0.12513445  3.2763408
## RubricTxtOrg                      1.01520604 0.13041244  7.7845799
## RubricVisOrg                      0.64658442 0.13476232  4.7979615
## as.factor(Rater)2:RubricInitEDA  -0.29999664 0.15672067 -1.9142124
## as.factor(Rater)3:RubricInitEDA  -0.29461500 0.15662493 -1.8810224
## as.factor(Rater)2:RubricInterpRes -0.51357660 0.15405307 -3.3337641
```

```
## as.factor(Rater)3:RubricInterpRes -0.71501679 0.15392098 -4.6453497
## as.factor(Rater)2:RubricRsrchQ    -0.48664032 0.14798221 -3.2885055
## as.factor(Rater)3:RubricRsrchQ    -0.32079180 0.14756365 -2.1739216
## as.factor(Rater)2:RubricSelMeth   -0.38584753 0.15092629 -2.5565296
## as.factor(Rater)3:RubricSelMeth   -0.38612008 0.14987370 -2.5763032
## as.factor(Rater)2:RubricTxtOrg    -0.55083463 0.15701123 -3.5082499
## as.factor(Rater)3:RubricTxtOrg    -0.44379962 0.15695213 -2.8276114
## as.factor(Rater)2:RubricVisOrg    -0.09926763 0.15966217 -0.6217355
## as.factor(Rater)3:RubricVisOrg    -0.26726239 0.15961681 -1.6744000
```

```
ranef(comb.final)
```

```
## $Artifact
##     RubricCritDes RubricInitEDA RubricInterpRes RubricRsrchQ RubricSelMeth
## 100   0.802136395  -0.261258264    -0.122829026 -0.166149960   0.230709645
## 101  -0.497513195   0.435020969    -0.167120800 -0.735163399   0.032862086
## 102  -0.766110424  -0.347220212    -0.242974352 -0.744974082   0.065147064
## 103   0.139624588   0.330146406     0.100448087 -0.284416550   0.268860217
## 104  -0.578097472   0.309048597     0.093152256 -0.262804860   0.072699210
## 105  -0.590884827  -0.480693475    -0.338646792 -0.402856992  -0.098806648
## 106   0.205004188  -1.028476514    -0.397542552  0.231924453  -0.136052892
## 107  -0.562041443  -0.400337174     0.057633850  0.182448959  -0.103670338
## 111  -0.463043057  -0.350097201    -0.240923801 -0.310649041  -0.068365403
## 112  -0.501881998   0.322315454     0.270722556  0.200618112  -0.102873082
## 113  -0.587810297  -0.802051471    -0.146639856 -0.128786943   0.002821561
## 114  -0.450255701   0.439644871     0.190875246 -0.170596909   0.103140456
## 115  -0.374040227   0.452911728     0.368445546  0.292826063  -0.072431837
## 116  -0.589706422  -0.352816959    -0.288269077 -0.345448181  -0.122388970
## 117  -0.673906733  -0.134262719     0.007010455 -0.182086280  -0.101964488
## 118  -0.655492622  -0.209995117    -0.031274140 -0.202825741  -0.026822428
## 13    0.387458176  -0.747800636    -0.434777205  0.045385225  -0.160543205
## 15    0.695090968   0.481334741    -0.047108295 -0.098360908  -0.048612691
## 16    0.685230953   1.153398215     0.315018524 -0.008664497   0.188257807
## 17    0.335529637  -0.088759611     0.066863388  0.547211293  -0.198167735
## 21    0.778118620   1.089153224     0.452972697  0.432762629   0.087412796
## 22    0.730838527   0.379657450     0.249729043  0.423882918   0.089682327
## 23   -0.274798565  -0.725034888    -0.327290914 -0.014262765  -0.177880785
## 24    0.049475596  -0.202510364    -0.149247682 -0.132267224   0.069191810
## 25    1.163666306   0.007037196    -0.288055919  0.162000680  -0.129599615
## 26   -0.865505548  -0.482840550    -0.159677763 -0.517999661   0.110398138
## 27    0.101119123   0.384770140     0.007854472 -0.176799183   0.091824205
## 28   -0.301014937  -0.592682089    -0.414281162 -0.407602137  -0.061160233
## 32    0.663360432   0.405268721     0.250508843  0.409234184   0.002857318
## 33   -0.084933985   0.149782013    -0.059097977 -0.454795876   0.166243654
## 34    0.466925521  -0.310258382    -0.061315809 -0.202921189   0.258189934
## 35   -0.743373903  -0.250737343    -0.086343290 -0.274624352  -0.096893862
## 36    0.049475596  -0.202510364    -0.149247682 -0.132267224   0.069191810
## 37    0.778403633  -0.157168305    -0.205770050 -0.024531481   0.102404931
## 38   -0.014445866  -0.205395284    -0.142865658 -0.166888609  -0.063680741
## 39   -0.531364614   0.246686898     0.207933236  0.136919548  -0.088281470
## 40    0.104675756   0.356273948     0.013456695 -0.196771833   0.045776663
## 45    0.015455627  -0.243393215    -0.171233231 -0.093584863   0.049686192
## 46    0.164978174   0.322679017    -0.014595917 -0.143299339   0.030853936
## 47    1.128680951  -0.182746630    -0.049330289  0.664810824  -0.151682083
## 48    0.534409515   0.654589624     0.227416690  0.121798178   0.239872170
```

```
## 49   -0.904212225   0.218783906     0.272759016  0.168749341  -0.186334633
## 53    1.144639085   0.106565267     0.017778946  0.237184424   0.116834213
## 54   -0.665795621   0.381493535    -0.090897741 -0.666263271   0.146645908
## 55   -0.148877627  -0.370095489     0.069319499  0.320753068  -0.133305967
## 56    0.691290629  -0.262007859    -0.276402792 -0.055779567  -0.061872926
## 57   -0.770092013  -0.324165072    -0.168937299 -0.251256720  -0.082244467
## 6    -0.675895807  -0.276348614    -0.087123090 -0.259975618  -0.010068853
## 61    0.002557723   0.332494434    -0.077296517 -0.171228752   0.018609165
## 62    1.390164923   0.999245596     0.303062196  0.487779106  -0.050740369
## 63    0.684116878   0.348208987     0.207088172  0.444305680  -0.018083244
## 64    0.553746691  -0.160360404    -0.077143400  0.057624369   0.063501344
## 65    0.836073693  -0.451135332    -0.408945279  0.255544081  -0.214697675
## 66    0.740952348   0.908663844     0.371042590  0.380980566  -0.041108997
## 67   -0.819758067   0.146255553     0.290722841  0.150566476  -0.189256213
## 68    0.633232230  -0.238296287     0.050566110  0.489384993  -0.040534937
## 7    -0.624252280   0.310931889     0.069979063 -0.304507577   0.012563541
## 72   -0.056788979   0.417635607     0.190387834 -0.069037579   0.021920187
## 73   -0.748424147  -0.931097865    -0.322360248 -0.187868668  -0.016662849
## 74   -1.051698592   0.087784256     0.115976457 -0.489881790   0.090988657
## 75   -0.044581236   0.335563119     0.147359057 -0.094253483   0.095584409
## 76    0.036127233  -0.326539193    -0.215854754 -0.144332305  -0.006764108
## 77   -0.168018397  -0.230514636    -0.012262119 -0.067482911  -0.014900596
## 78    0.416923176   0.062461452     0.075227870  0.130322947   0.083909720
## 79   -0.310792734   0.018357636     0.132676795  0.023461469   0.050985226
## 8    -0.609224512  -0.208240807    -0.036159166 -0.211887865   0.005806688
## 84    0.307626689  -0.084549887     0.158170961  0.448254647  -0.060807587
## 85    1.074043498   0.374478907     0.314881074  0.875679474  -0.131295658
## 86    0.326040799  -0.160282285     0.119886366  0.427515186   0.014334473
## 87    0.205004188  -1.028476514    -0.397542552  0.231924453  -0.136052892
## 88    1.088841574   0.643273121     0.311602954  0.545943375  -0.075566241
## 9    -0.583008141  -0.340593605     0.050831083  0.181451507  -0.110913865
## 92   -0.542448927  -0.348193057     0.069726970  0.219118309  -0.052110600
## 93   -0.395014641  -0.165452666     0.179543080  0.347995610   0.029890383
## 94    1.038394953   1.031720561     0.338306360  0.393756305  -0.006485848
## 95    0.139624588   0.330146406     0.100448087 -0.284416550   0.268860217
## 96    0.238765958   0.379629951     0.223158962  0.316807038  -0.066869254
## O1   -0.502091278   0.422490469     0.123669006 -0.011803265  -0.093732258
## O10  -0.442591668  -0.001088562     0.155367704  0.049395438   0.036686599
## O11  -0.768621200  -0.331059930     0.329692036  0.509151774   0.013836581
## O12  -0.354612131  -0.487432726    -0.314831682 -0.339630738  -0.017019066
## O13   0.031532246   0.083791406    -0.315396900 -0.366236173  -0.069771388
## O2   -0.117227703   0.329710464     0.169735305  0.140797495  -0.070993748
## O3    0.284397693  -0.134149047    -0.012522254  0.231842620  -0.038358577
## O4   -0.111505301   0.045232807     0.204675572 -0.217979623   0.362516806
## O5    0.879664007  -0.426528072    -0.024422086  0.186630268   0.246928845
## O6   -0.897859646  -0.771486533    -0.418455243 -0.410714553  -0.110919937
## O7    0.138404024   0.207476794    -0.005323983 -0.012193932  -0.041817221
## O8    0.358541891  -0.211383044    -0.392531602 -0.318254832   0.024074883
## O9   -0.800750292   0.586409075     0.347184395 -0.187357233   0.076606704
##     RubricTxtOrg RubricVisOrg as.factor(Rater)1 as.factor(Rater)2
## 100 -0.489209779  -0.46338654       0.035228852      -0.050459504
## 101  0.294213529   0.46924474      -0.031810421       0.045563167
## 102 -1.138166812  -0.58784053      -0.076293659       0.109278048
## 103  0.107433142  -0.24044501       0.042808676      -0.061316348
```

```
## 104  0.086517991  -0.11636154      -0.025034756       0.035858148
## 105 -0.529793298  -0.35679039      -0.059307693       0.084948461
## 106  0.015382372  -0.32465661      -0.022821791       0.032688440
## 107 -0.512120155  -0.30914586      -0.010612037       0.015199988
## 111 -0.409135219  -0.25242074      -0.036045768       0.051629601
## 112  0.210684913   0.42590095       0.019639505      -0.028130342
## 113 -0.469075365  -0.30975769      -0.016156619       0.023141684
## 114  0.207176070  -0.01199188      -0.001772831       0.002539287
## 115  0.331342992   0.53027060       0.042901430      -0.061449203
## 116  0.134854799   0.28665626      -0.031846463       0.045614790
## 117  0.237782226   0.84980061       0.009491064      -0.013594379
## 118  0.133071623   0.32091946      -0.009167002       0.013130214
## 13  -0.724907116  -0.63443287      -0.065672904      -0.384872005
## 15   0.423411083   0.91043882       0.013328123       0.078108646
## 16   0.896257617   0.58846405       0.020598052       0.120713614
## 17  -0.117441051   0.03320250      -0.018086663      -0.105995773
## 21   0.914069901   0.60042329       0.043540823       0.255168310
## 22   0.203420469  -0.12410885       0.018626171       0.109157527
## 23  -0.674859692  -0.56622161      -0.056917691      -0.333562619
## 24   0.242494595  -0.13904091      -0.007419533      -0.043481716
## 25  -0.001380211   0.07631826      -0.038683242      -0.226700755
## 26  -0.468813235  -0.47687729       0.015918252       0.093287933
## 27   0.293726362  -0.05343108      -0.006120428      -0.035868388
## 28  -0.629747077  -0.52375060      -0.069030239      -0.404547461
## 32   0.261824093   0.36471660       0.027312942       0.160065811
## 33  -0.404703994  -0.40504847       0.019292367       0.113061728
## 34  -0.501129039  -0.51886505       0.030627312       0.179489473
## 35  -0.250068398   0.26789895      -0.004906325      -0.028753218
## 36   0.242494595  -0.13904091      -0.007419533      -0.043481716
## 37   0.255372069  -0.15282192      -0.005344523      -0.031321249
## 38  -0.237190924   0.25411794      -0.002831316      -0.016592751
## 39  -0.239427971  -0.12335744       0.010648781       0.062406522
## 40  -0.244362781  -0.14909769      -0.010218981      -0.059887706
## 45   0.293741321  -0.21288497       0.014005375      -0.084284952
## 46  -0.190078179  -0.22348797       0.017686876      -0.106440383
## 47  -0.693767136  -0.67950859       0.060465786      -0.363885714
## 48   0.594297202  -0.35248561      -0.059095089       0.355636795
## 49   0.268325659   0.73774352      -0.027603923       0.166121598
## 53   0.073299989  -0.06127576      -0.066884744       0.402515276
## 54   0.328570477  -0.11566550       0.048138267      -0.289698170
## 55  -0.359900249   0.05958794      -0.010164118       0.061168099
## 56  -0.232483900   0.11252625       0.019745764      -0.118830861
## 57   0.279878670   0.23577941       0.019820502      -0.119280639
## 6   -0.308472022  -0.22092651      -0.013593096      -0.079661501
## 61   0.877957149   0.39971111       0.008968864      -0.053975010
## 62   0.410995441   0.82768734      -0.054257702       0.326525194
## 63   0.295132103   0.27278493      -0.036743159       0.221121916
## 64   0.239453322   0.18677687      -0.001539659       0.009265733
## 65   0.315447231   0.17745210       0.011537970      -0.069436003
## 66  -0.190181855   0.26406321      -0.032677865       0.196656802
## 67  -0.856947666   0.06563681      -0.003085542       0.018568925
## 68   0.244616571   0.18779465      -0.034856627       0.209768688
## 7   -0.257240255  -0.13531668      -0.012293990      -0.072048173
## 72  -0.200640221   0.24209103      -0.010188485       0.061314746
```

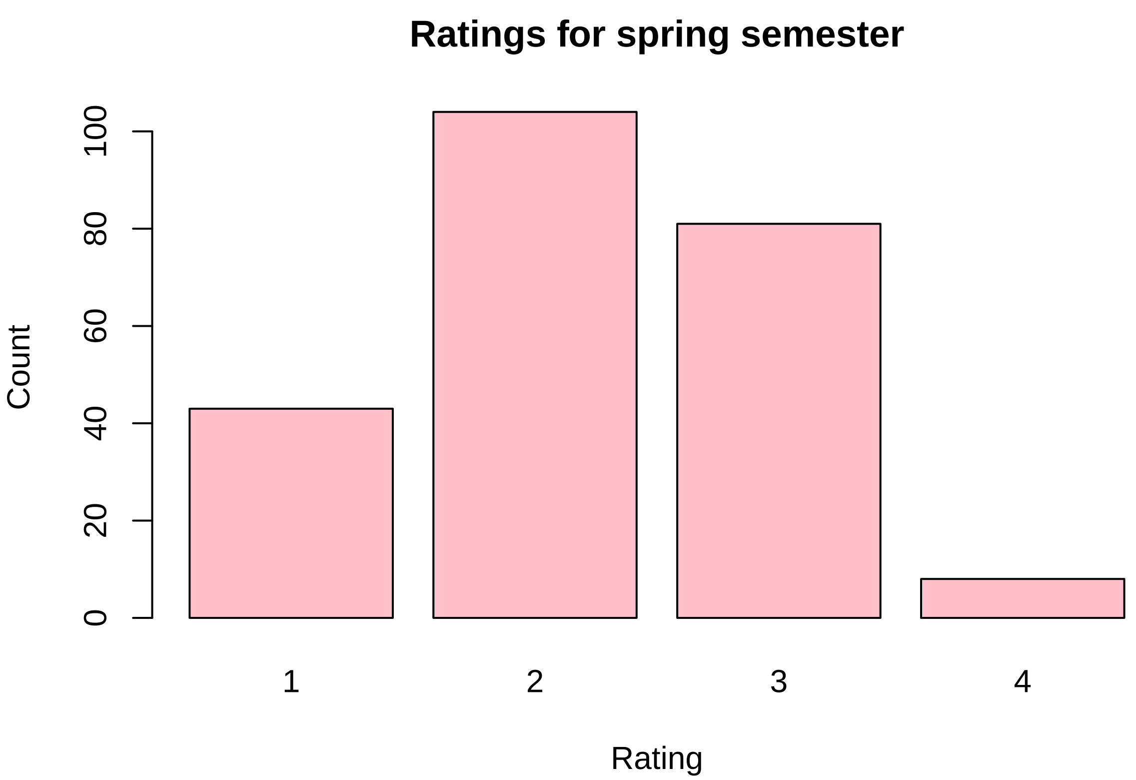
```
## 73   0.175748791  -0.33389358      0.034321764     -0.206549857
## 74  -0.445939253  -0.06242849     -0.034303507      0.206439986
## 75  -0.311208443  -0.29185698      0.018153647     -0.109249431
## 76  -0.297831707  -0.35738588      0.035373776     -0.212880911
## 77  -0.308109626   0.10783546      0.007425448     -0.044686669
## 78   0.061566530  -0.04650518     -0.063871462      0.384381216
## 79   0.049833071  -0.03173459     -0.060858180      0.366247156
## 8   -0.245547124  -0.16649625     -0.002762873     -0.016191648
## 84   0.297124849   0.43445755      0.044567972     -0.063836249
## 85   0.278018885   0.43085648      0.054569408     -0.078161652
## 86   0.192414246  -0.09442360      0.025909907     -0.037111656
## 87   0.015382372  -0.32465661     -0.022821791      0.032688440
## 88   0.483931248   1.05072595      0.070473204     -0.100941209
## 9   -0.290659738  -0.20896727      0.009349675      0.054793194
## 92   0.047817340  -0.19458036     -0.001808873      0.002590911
## 93   0.728412913   0.02435480      0.030256217     -0.043337027
## 94   0.316719536   0.50907530      0.031343524     -0.044894415
## 95   0.107433142  -0.24044501      0.042808676     -0.061316348
## 96   0.234623739   0.42168805      0.024096359     -0.034514049
## O1  -0.238697410  -0.25934972     -0.212788544      0.271957829
## O10  0.120887893   0.01366580      0.109293378     -0.368007716
## O11  0.301689020  -0.25762889      0.053178789      0.052114806
## O12 -0.363216422  -0.45575899      0.058776851     -0.016566939
## O13  0.274972868   0.15304202     -0.008262062      0.047391105
## O2   0.173695970   0.34400302      0.173370549     -1.028395017
## O3   0.243500276   0.13620567     -0.038664758      0.158496247
## O4   0.161554727  -0.27692379     -0.089484223      0.417972227
## O5  -0.043206811  -0.48792808      0.031262708      0.059900622
## O6  -0.053722591  -0.17911064     -0.050366154      0.114828728
## O7   0.124255514   0.22882082      0.148206776      0.138024125
## O8  -0.572731138  -0.91364174     -0.038888128     -0.247306895
## O9   0.399285360   0.56141963      0.048314585      0.033962466
##     as.factor(Rater)3
## 100       0.031450251
## 101      -0.028398476
## 102      -0.068110500
## 103       0.038217073
## 104      -0.022349561
## 105      -0.052946427
## 106      -0.020373955
## 107      -0.009473804
## 111      -0.032179545
## 112       0.017532997
## 113      -0.014423681
## 114      -0.001582679
## 115       0.038299878
## 116      -0.028430652
## 117       0.008473064
## 118      -0.008183761
## 13       -0.533359241
## 15        0.108243696
## 16        0.167286061
## 17       -0.146889939
## 21        0.353614642
```

```
## 22        0.151271526
## 23       -0.462254211
## 24       -0.060257371
## 25       -0.314164036
## 26        0.129279294
## 27       -0.049706750
## 28       -0.560625672
## 32        0.221820704
## 33        0.156682005
## 34        0.248738198
## 35       -0.039846479
## 36       -0.060257371
## 37       -0.043405282
## 38       -0.022994390
## 39        0.086483545
## 40       -0.082992946
## 45       -0.051157108
## 46       -0.064604440
## 47       -0.220861972
## 48        0.215855256
## 49        0.100828206
## 53        0.244308348
## 54       -0.175833529
## 55        0.037126236
## 56       -0.072124893
## 57       -0.072397888
## 6        -0.110395657
## 61       -0.032760360
## 62        0.198185847
## 63        0.134210881
## 64        0.005623876
## 65       -0.042144476
## 66        0.119361677
## 67        0.011270488
## 68        0.127319992
## 7        -0.099845036
## 72        0.037215244
## 73       -0.125366309
## 74        0.125299622
## 75       -0.066309404
## 76       -0.129208969
## 77       -0.027122762
## 78        0.233301804
## 79        0.222295259
## 8        -0.022438538
## 84        0.039787669
## 85        0.048716364
## 86        0.023130844
## 87       -0.020373955
## 88        0.062914339
## 9         0.075932923
## 92       -0.001614855
## 93        0.027010975
## 94        0.027981659
```

```
## 95        0.038217073
## 96        0.021511814
## 01       -0.222503905
## 010      -0.112034300
## 011       0.174632715
## 012       0.119500282
## 013       0.027869013
## 02       -0.618441949
## 03        0.067691812
## 04        0.207368016
## 05        0.131669045
## 06       -0.002651372
## 07        0.479539506
## 08       -0.335062767
## 09        0.145391452
##
## with conditional variances for "Artifact"
```

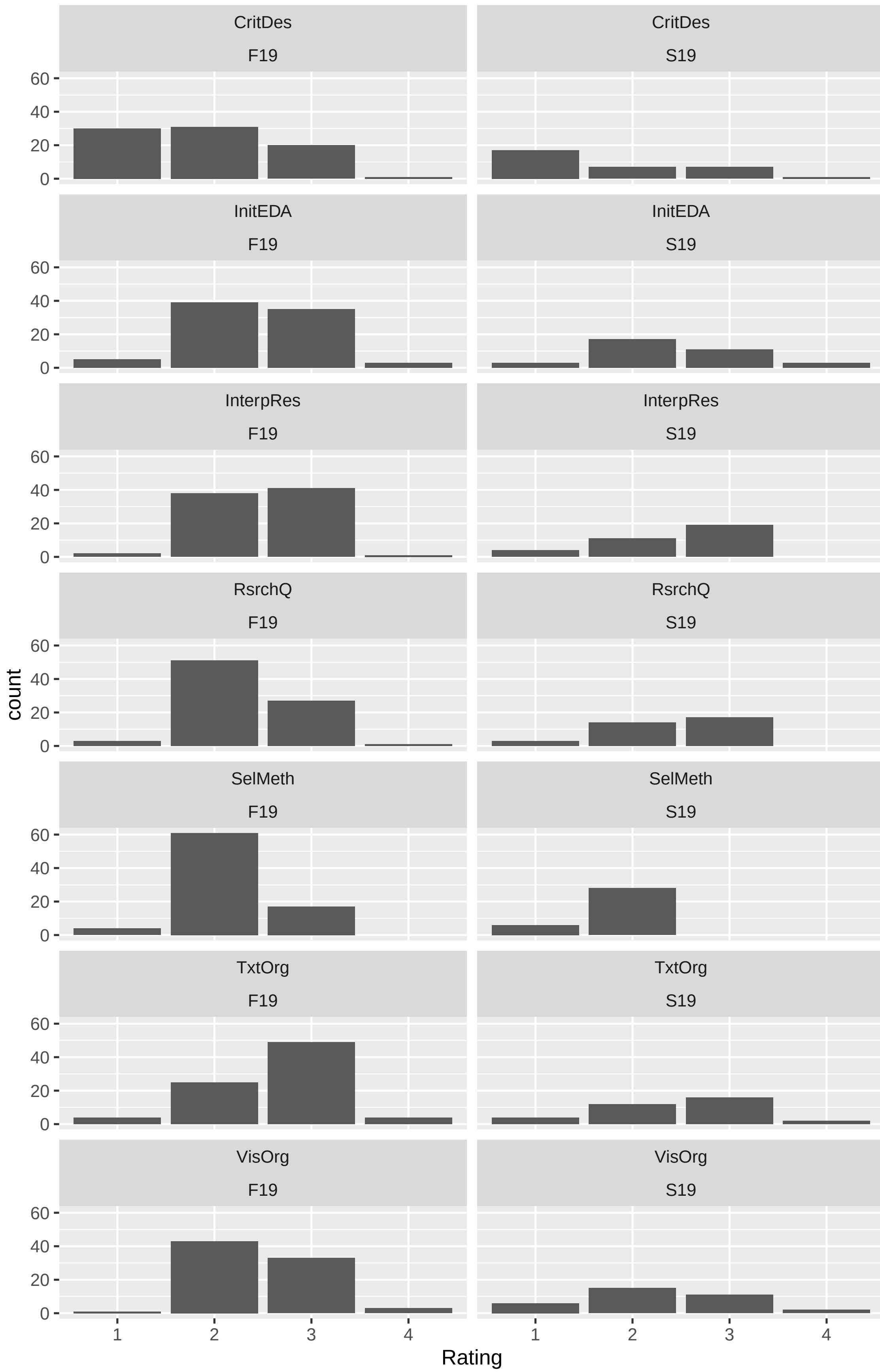
Part 4

```
spring <- tall.nonmissing[tall.nonmissing$Semester=="S19",]
fall <- tall.nonmissing[tall.nonmissing$Semester=="F19",]
barplot(table(spring$Rating), main = "Ratings for spring semester", xlab = "Rating", ylab = "Count", col
```




```
barplot(table(fall$Rating), main = "Ratings for fall semester", xlab = "Rating", ylab = "Count", col =
```


```
g <- ggplot(tall.nonmissing, aes(x=Rating)) +
  geom_bar() +
  facet_wrap( ~ Rubric + Semester, nrow=7)
g
```


CritDes
F19
S19
InitEDA
InterpRes
RsrchQ
SelMeth
TxtOrg
VisOrg
count
60
40
20
0
1
2
3
4
Rating

```r
summary(spring)
```

```
##        X             Rater         Artifact            Repeated     
##  Min.   :  3.0   Min.   :1.00   Length:236         Min.   :0.0000  
##  1st Qu.:206.2   1st Qu.:1.00   Class :character   1st Qu.:0.0000  
##  Median :412.5   Median :2.00   Mode  :character   Median :0.0000  
##  Mean   :409.8   Mean   :1.97                      Mean   :0.2669  
##  3rd Qu.:609.8   3rd Qu.:3.00                      3rd Qu.:1.0000  
##  Max.   :816.0   Max.   :3.00                      Max.   :1.0000  
##    Semester             Sex               Rubric          Rating 
##  Length:236         Length:236         Length:236         1: 43  
##  Class :character   Class :character   Class :character   2:104  
##  Mode  :character   Mode  :character   Mode  :character   3: 81  
##                                                           4:  8  
## 
## 
```

```r
summary(fall)
```

```
##        X             Rater          Artifact            Repeated     
##  Min.   :  1.0   Min.   :1.000   Length:572         Min.   :0.0000  
##  1st Qu.:207.5   1st Qu.:1.000   Class :character   1st Qu.:0.0000  
##  Median :409.5   Median :2.000   Mode  :character   Median :0.0000  
##  Mean   :410.6   Mean   :2.003                      Mean   :0.3671  
##  3rd Qu.:615.2   3rd Qu.:3.000                      3rd Qu.:1.0000  
##  Max.   :819.0   Max.   :3.000                      Max.   :1.0000  
##    Semester             Sex               Rubric          Rating 
##  Length:572         Length:572         Length:572         1: 49  
##  Class :character   Class :character   Class :character   2:288  
##  Mode  :character   Mode  :character   Mode  :character   3:222  
##                                                           4: 13  
## 
## 
```

```r
43/236
```

```
## [1] 0.1822034
```

```r
104/236
```

```
## [1] 0.440678
```

```r
81/236
```

```
## [1] 0.3432203
```

```r
8/236
```

```
## [1] 0.03389831
```

```r
49/572
```

```
## [1] 0.08566434
```

```r
288/572
```

```
## [1] 0.5034965
```

```r
222/572
```

```
## [1] 0.3881119
```

```r
13/572
```

```
## [1] 0.02272727
```

```r
s_r <- spring[spring$Rubric=="RsrchQ",]
summary(s_r)
```

```
##        X              Rater          Artifact            Repeated     
##  Min.   :  3.00   Min.   :1.000   Length:34          Min.   :0.0000  
##  1st Qu.: 24.75   1st Qu.:1.000   Class :character   1st Qu.:0.0000  
##  Median : 56.00   Median :2.000   Mode  :character   Median :0.0000  
##  Mean   : 56.74   Mean   :1.971                      Mean   :0.2647  
##  3rd Qu.: 84.50   3rd Qu.:3.000                      3rd Qu.:0.7500  
##  Max.   :114.00   Max.   :3.000                      Max.   :1.0000  
##    Semester             Sex               Rubric          Rating
##  Length:34          Length:34          Length:34          1: 3  
##  Class :character   Class :character   Class :character   2:14  
##  Mode  :character   Mode  :character   Mode  :character   3:17  
##                                                           4: 0  
## 
## 
```

```r
3/34
```

```
## [1] 0.08823529
```

```r
14/34
```

```
## [1] 0.4117647
```

```r
17/34
```

```
## [1] 0.5
```

```r
0
```

```
## [1] 0
```

```r
s_init <- spring[spring$Rubric=="InitEDA",]
summary(s_init)
```

```
##        X             Rater          Artifact            Repeated     
##  Min.   :237.0   Min.   :1.000   Length:34          Min.   :0.0000  
##  1st Qu.:258.8   1st Qu.:1.000   Class :character   1st Qu.:0.0000  
##  Median :290.0   Median :2.000   Mode  :character   Median :0.0000  
##  Mean   :290.7   Mean   :1.971                      Mean   :0.2647  
##  3rd Qu.:318.5   3rd Qu.:3.000                      3rd Qu.:0.7500  
##  Max.   :348.0   Max.   :3.000                      Max.   :1.0000  
##    Semester             Sex               Rubric          Rating
##  Length:34          Length:34          Length:34          1: 3  
##  Class :character   Class :character   Class :character   2:17  
##  Mode  :character   Mode  :character   Mode  :character   3:11  
##                                                           4: 3  
## 
## 
```

```r
3/34
```

```
## [1] 0.08823529
```

```r
17/34
```

```
## [1] 0.5
```

```r
11/34
```

```
## [1] 0.3235294
```

```r
3/34
```

```
## [1] 0.08823529
```

```r
s_inter <- spring[spring$Rubric=="InterpRes",]
summary(s_inter)
```

```
##        X              Rater         Artifact            Repeated
##  Min.   :471.0   Min.   :1.000   Length:34          Min.   :0.0000
##  1st Qu.:492.8   1st Qu.:1.000   Class :character   1st Qu.:0.0000
##  Median :524.0   Median :2.000   Mode  :character   Median :0.0000
##  Mean   :524.7   Mean   :1.971                      Mean   :0.2647
##  3rd Qu.:552.5   3rd Qu.:3.000                      3rd Qu.:0.7500
##  Max.   :582.0   Max.   :3.000                      Max.   :1.0000
##    Semester             Sex               Rubric          Rating
##  Length:34          Length:34          Length:34          1: 4
##  Class :character   Class :character   Class :character   2:11
##  Mode  :character   Mode  :character   Mode  :character   3:19
##                                                           4: 0
##
##
```

```r
4/34
```

```
## [1] 0.1176471
```

```r
11/34
```

```
## [1] 0.3235294
```

```r
19/34
```

```
## [1] 0.5588235
```

```r
0
```

```
## [1] 0
```

```r
s_crit <- spring[spring$Rubric=="CritDes",]
summary(s_crit)
```

```
##        X              Rater         Artifact            Repeated
##  Min.   :120.0   Min.   :1.000   Length:32          Min.   :0.0000
##  1st Qu.:140.5   1st Qu.:1.000   Class :character   1st Qu.:0.0000
##  Median :178.5   Median :2.000   Mode  :character   Median :0.0000
##  Mean   :174.5   Mean   :1.969                      Mean   :0.2812
##  3rd Qu.:202.5   3rd Qu.:3.000                      3rd Qu.:1.0000
##  Max.   :231.0   Max.   :3.000                      Max.   :1.0000
##    Semester             Sex               Rubric          Rating
##  Length:32          Length:32          Length:32          1:17
##  Class :character   Class :character   Class :character   2: 7
##  Mode  :character   Mode  :character   Mode  :character   3: 7
```

```
##                                                             4: 1
##
##
```

```r
17/32
```

```
## [1] 0.53125
```

```r
8/32
```

```
## [1] 0.25
```

```r
7/32
```

```
## [1] 0.21875
```

```r
1/32
```

```
## [1] 0.03125
```

```r
s_sel <- spring[spring$Rubric=="SelMeth",]
summary(s_sel)
```

```
##        X             Rater          Artifact            Repeated
##  Min.   :354.0   Min.   :1.000   Length:34          Min.   :0.0000
##  1st Qu.:375.8   1st Qu.:1.000   Class :character   1st Qu.:0.0000
##  Median :407.0   Median :2.000   Mode  :character   Median :0.0000
##  Mean   :407.7   Mean   :1.971                      Mean   :0.2647
##  3rd Qu.:435.5   3rd Qu.:3.000                      3rd Qu.:0.7500
##  Max.   :465.0   Max.   :3.000                      Max.   :1.0000
##    Semester             Sex               Rubric           Rating
##  Length:34          Length:34          Length:34          1: 6
##  Class :character   Class :character   Class :character   2:28
##  Mode  :character   Mode  :character   Mode  :character   3: 0
##                                                           4: 0
##
##
```

```r
6/34
```

```
## [1] 0.1764706
```

```r
28/34
```

```
## [1] 0.8235294
```

```r
s_txt <- spring[spring$Rubric=="TxtOrg",]
summary(s_txt)
```

```
##        X             Rater          Artifact            Repeated
##  Min.   :705.0   Min.   :1.000   Length:34          Min.   :0.0000
##  1st Qu.:726.8   1st Qu.:1.000   Class :character   1st Qu.:0.0000
##  Median :758.0   Median :2.000   Mode  :character   Median :0.0000
##  Mean   :758.7   Mean   :1.971                      Mean   :0.2647
##  3rd Qu.:786.5   3rd Qu.:3.000                      3rd Qu.:0.7500
##  Max.   :816.0   Max.   :3.000                      Max.   :1.0000
##    Semester             Sex               Rubric           Rating
##  Length:34          Length:34          Length:34          1: 4
##  Class :character   Class :character   Class :character   2:12
##  Mode  :character   Mode  :character   Mode  :character   3:16
```

```
##                                                              4: 2
##
##
```

```r
4/34
```

```
## [1] 0.1176471
```

```r
12/34
```

```
## [1] 0.3529412
```

```r
16/34
```

```
## [1] 0.4705882
```

```r
2/34
```

```
## [1] 0.05882353
```

```r
s_vis <- spring[spring$Rubric=="VisOrg",]
summary(s_vis)
```

```
##        X              Rater          Artifact            Repeated
##  Min.   :588.0   Min.   :1.000   Length:34          Min.   :0.0000
##  1st Qu.:609.8   1st Qu.:1.000   Class :character   1st Qu.:0.0000
##  Median :641.0   Median :2.000   Mode  :character   Median :0.0000
##  Mean   :641.7   Mean   :1.971                      Mean   :0.2647
##  3rd Qu.:669.5   3rd Qu.:3.000                      3rd Qu.:0.7500
##  Max.   :699.0   Max.   :3.000                      Max.   :1.0000
##    Semester             Sex               Rubric           Rating
##  Length:34          Length:34          Length:34          1: 6
##  Class :character   Class :character   Class :character   2:15
##  Mode  :character   Mode  :character   Mode  :character   3:11
##                                                           4: 2
##
##
```

```r
6/34
```

```
## [1] 0.1764706
```

```r
15/34
```

```
## [1] 0.4411765
```

```r
11/34
```

```
## [1] 0.3235294
```

```r
2/34
```

```
## [1] 0.05882353
```

```r
f_r <- fall[fall$Rubric=="RsrchQ",]
summary(f_r)
```

```
##        X               Rater     Artifact           Repeated
##  Min.   :  1.00   Min.   :1   Length:82          Min.   :0.0000
##  1st Qu.: 32.25   1st Qu.:1   Class :character   1st Qu.:0.0000
##  Median : 59.50   Median :2   Mode  :character   Median :0.0000
##  Mean   : 60.60   Mean   :2                      Mean   :0.3659
```

```
##  3rd Qu.: 90.50   3rd Qu.:3                      3rd Qu.:1.0000
##  Max.   :117.00   Max.   :3                      Max.   :1.0000
##    Semester             Sex                Rubric            Rating
##  Length:82          Length:82          Length:82          1: 3
##  Class :character   Class :character   Class :character   2:51
##  Mode  :character   Mode  :character   Mode  :character   3:27
##                                                           4: 1
##
##
```

```r
3/83
```

```
## [1] 0.03614458
```

```r
51/83
```

```
## [1] 0.6144578
```

```r
28/83
```

```
## [1] 0.3373494
```

```r
1/83
```

```
## [1] 0.01204819
```

```r
f_init <- fall[fall$Rubric=="InitEDA",]
summary(f_init)
```

```
##        X             Rater     Artifact            Repeated
##  Min.   :235.0   Min.   :1   Length:82          Min.   :0.0000
##  1st Qu.:266.2   1st Qu.:1   Class :character   1st Qu.:0.0000
##  Median :293.5   Median :2   Mode  :character   Median :0.0000
##  Mean   :294.6   Mean   :2                      Mean   :0.3659
##  3rd Qu.:324.5   3rd Qu.:3                      3rd Qu.:1.0000
##  Max.   :351.0   Max.   :3                      Max.   :1.0000
##    Semester             Sex                Rubric            Rating
##  Length:82          Length:82          Length:82          1: 5
##  Class :character   Class :character   Class :character   2:39
##  Mode  :character   Mode  :character   Mode  :character   3:35
##                                                           4: 3
##
##
```

```r
5/82
```

```
## [1] 0.06097561
```

```r
39/82
```

```
## [1] 0.4756098
```

```r
35/82
```

```
## [1] 0.4268293
```

```r
3/82
```

```
## [1] 0.03658537
```

```r
f_inter <- fall[fall$Rubric=="InterpRes",]
summary(f_inter)
```

```
##        X             Rater     Artifact           Repeated
##  Min.   :469.0   Min.   :1   Length:82          Min.   :0.0000
##  1st Qu.:500.2   1st Qu.:1   Class :character   1st Qu.:0.0000
##  Median :527.5   Median :2   Mode  :character   Median :0.0000
##  Mean   :528.6   Mean   :2                      Mean   :0.3659
##  3rd Qu.:558.5   3rd Qu.:3                      3rd Qu.:1.0000
##  Max.   :585.0   Max.   :3                      Max.   :1.0000
##    Semester             Sex               Rubric           Rating
##  Length:82          Length:82          Length:82          1: 2
##  Class :character   Class :character   Class :character   2:38
##  Mode  :character   Mode  :character   Mode  :character   3:41
##                                                           4: 1
##
##
```

```
2/82
```

```
## [1] 0.02439024
```

```
38/82
```

```
## [1] 0.4634146
```

```
41/82
```

```
## [1] 0.5
```

```
1/82
```

```
## [1] 0.01219512
```

```
f_crit <- fall[fall$Rubric=="CritDes",]
summary(f_crit)
```

```
##        X             Rater     Artifact           Repeated
##  Min.   :118.0   Min.   :1   Length:82          Min.   :0.0000
##  1st Qu.:149.2   1st Qu.:1   Class :character   1st Qu.:0.0000
##  Median :176.5   Median :2   Mode  :character   Median :0.0000
##  Mean   :177.6   Mean   :2                      Mean   :0.3659
##  3rd Qu.:207.5   3rd Qu.:3                      3rd Qu.:1.0000
##  Max.   :234.0   Max.   :3                      Max.   :1.0000
##    Semester             Sex               Rubric           Rating
##  Length:82          Length:82          Length:82          1:30
##  Class :character   Class :character   Class :character   2:31
##  Mode  :character   Mode  :character   Mode  :character   3:20
##                                                           4: 1
##
##
```

```
30/82
```

```
## [1] 0.3658537
```

```
31/82
```

```
## [1] 0.3780488
```

```
20/82
```

```
## [1] 0.2439024
```

```r
1/82
```

```
## [1] 0.01219512
```

```r
f_sel <- fall[fall$Rubric=="SelMeth",]
summary(f_sel)
```

```
##        X              Rater     Artifact            Repeated
##  Min.   :352.0   Min.   :1   Length:82          Min.   :0.0000
##  1st Qu.:383.2   1st Qu.:1   Class :character   1st Qu.:0.0000
##  Median :410.5   Median :2   Mode  :character   Median :0.0000
##  Mean   :411.6   Mean   :2                      Mean   :0.3659
##  3rd Qu.:441.5   3rd Qu.:3                      3rd Qu.:1.0000
##  Max.   :468.0   Max.   :3                      Max.   :1.0000
##    Semester             Sex               Rubric           Rating
##  Length:82          Length:82          Length:82          1: 4
##  Class :character   Class :character   Class :character   2:61
##  Mode  :character   Mode  :character   Mode  :character   3:17
##                                                           4: 0
##
##
```

```r
4/82
```

```
## [1] 0.04878049
```

```r
61/82
```

```
## [1] 0.7439024
```

```r
17/82
```

```
## [1] 0.2073171
```

```r
f_txt <- fall[fall$Rubric=="TxtOrg",]
summary(f_txt)
```

```
##        X              Rater     Artifact            Repeated
##  Min.   :703.0   Min.   :1   Length:82          Min.   :0.0000
##  1st Qu.:734.2   1st Qu.:1   Class :character   1st Qu.:0.0000
##  Median :761.5   Median :2   Mode  :character   Median :0.0000
##  Mean   :762.6   Mean   :2                      Mean   :0.3659
##  3rd Qu.:792.5   3rd Qu.:3                      3rd Qu.:1.0000
##  Max.   :819.0   Max.   :3                      Max.   :1.0000
##    Semester             Sex               Rubric           Rating
##  Length:82          Length:82          Length:82          1: 4
##  Class :character   Class :character   Class :character   2:25
##  Mode  :character   Mode  :character   Mode  :character   3:49
##                                                           4: 4
##
##
```

```r
4/82
```

```
## [1] 0.04878049
```

```r
25/82
```

```
## [1] 0.304878
```

```r
49/82
```

```
## [1] 0.597561
```

```r
4/82
```

```
## [1] 0.04878049
```

```r
f_vis <- fall[fall$Rubric=="VisOrg",]
summary(f_vis)
```

```
##        X             Rater         Artifact           Repeated
##  Min.   :586.0   Min.   :1.000   Length:80          Min.   :0.000
##  1st Qu.:616.8   1st Qu.:1.000   Class :character   1st Qu.:0.000
##  Median :643.5   Median :2.000   Mode  :character   Median :0.000
##  Mean   :644.6   Mean   :2.025                      Mean   :0.375
##  3rd Qu.:673.2   3rd Qu.:3.000                      3rd Qu.:1.000
##  Max.   :702.0   Max.   :3.000                      Max.   :1.000
##    Semester             Sex                Rubric           Rating
##  Length:80          Length:80          Length:80          1: 1
##  Class :character   Class :character   Class :character   2:43
##  Mode  :character   Mode  :character   Mode  :character   3:33
##                                                           4: 3
##
##
```

```r
1/80
```

```
## [1] 0.0125
```

```r
43/80
```

```
## [1] 0.5375
```

```r
33/80
```

```
## [1] 0.4125
```

```r
3/80
```

```
## [1] 0.0375
```

```r
sr1 <- spring[spring$Rater==1,]
summary(sr1)
```

```
##        X             Rater     Artifact           Repeated
##  Min.   : 81.0   Min.   :1   Length:84          Min.   :0.00
##  1st Qu.:228.5   1st Qu.:1   Class :character   1st Qu.:0.00
##  Median :442.0   Median :1   Mode  :character   Median :0.00
##  Mean   :446.8   Mean   :1                      Mean   :0.25
##  3rd Qu.:668.5   3rd Qu.:1                      3rd Qu.:0.25
##  Max.   :816.0   Max.   :1                      Max.   :1.00
##    Semester             Sex                Rubric           Rating
##  Length:84          Length:84          Length:84          1:11
##  Class :character   Class :character   Class :character   2:38
##  Mode  :character   Mode  :character   Mode  :character   3:33
##                                                           4: 2
##
##
```

```r
11/84
```

```
## [1] 0.1309524
```

```r
38/84
```

```
## [1] 0.452381
```

```r
33/84
```

```
## [1] 0.3928571
```

```r
2/84
```

```
## [1] 0.02380952
```

```r
sr2 <- spring[spring$Rater==2,]
summary(sr2)
```

```
##        X             Rater     Artifact            Repeated   
##  Min.   : 42.0   Min.   :2   Length:75          Min.   :0.00  
##  1st Qu.:186.0   1st Qu.:2   Class :character   1st Qu.:0.00  
##  Median :412.0   Median :2   Mode  :character   Median :0.00  
##  Mean   :411.8   Mean   :2                      Mean   :0.28  
##  3rd Qu.:629.5   3rd Qu.:2                      3rd Qu.:1.00  
##  Max.   :773.0   Max.   :2                      Max.   :1.00  
##    Semester             Sex                Rubric           Rating
##  Length:75          Length:75          Length:75          1: 7  
##  Class :character   Class :character   Class :character   2:32  
##  Mode  :character   Mode  :character   Mode  :character   3:32  
##                                                           4: 4  
## 
## 
```

```r
7/75
```

```
## [1] 0.09333333
```

```r
32/75
```

```
## [1] 0.4266667
```

```r
32/75
```

```
## [1] 0.4266667
```

```r
4/75
```

```
## [1] 0.05333333
```

```r
sr3 <- spring[spring$Rater==3,]
summary(sr3)
```

```
##        X             Rater     Artifact            Repeated     
##  Min.   :  3.0   Min.   :3   Length:77          Min.   :0.0000  
##  1st Qu.:141.0   1st Qu.:3   Class :character   1st Qu.:0.0000  
##  Median :367.0   Median :3   Mode  :character   Median :0.0000  
##  Mean   :367.6   Mean   :3                      Mean   :0.2727  
##  3rd Qu.:593.0   3rd Qu.:3                      3rd Qu.:1.0000  
##  Max.   :740.0   Max.   :3                      Max.   :1.0000  
##    Semester             Sex                Rubric           Rating
```

```
##   Length:77          Length:77          Length:77          1:25
##   Class :character   Class :character   Class :character   2:34
##   Mode  :character   Mode  :character   Mode  :character   3:16
##                                                            4: 2
##
##
```

```r
25/77
```

```
## [1] 0.3246753
```

```r
34/77
```

```
## [1] 0.4415584
```

```r
16/77
```

```
## [1] 0.2077922
```

```r
2/77
```

```
## [1] 0.02597403
```

```r
fr1 <- fall[fall$Rater==1,]
summary(fr1)
```

```
##         X             Rater     Artifact            Repeated
##   Min.   : 79.0   Min.   :1   Length:187         Min.   :0.0000
##   1st Qu.:223.5   1st Qu.:1   Class :character   1st Qu.:0.0000
##   Median :449.0   Median :1   Mode  :character   Median :0.0000
##   Mean   :447.5   Mean   :1                      Mean   :0.3743
##   3rd Qu.:673.5   3rd Qu.:1                      3rd Qu.:1.0000
##   Max.   :819.0   Max.   :1                      Max.   :1.0000
##     Semester             Sex                Rubric           Rating
##   Length:187         Length:187         Length:187         1:18
##   Class :character   Class :character   Class :character   2:86
##   Mode  :character   Mode  :character   Mode  :character   3:79
##                                                            4: 4
##
##
```

```r
18/187
```

```
## [1] 0.09625668
```

```r
86/187
```

```
## [1] 0.459893
```

```r
79/187
```

```
## [1] 0.4224599
```

```r
4/187
```

```
## [1] 0.02139037
```

```r
fr2 <- fall[fall$Rater==2,]
summary(fr2)
```

```
##         X             Rater     Artifact            Repeated
##   Min.   : 40.0   Min.   :2   Length:196         Min.   :0.0000
```

```
##  1st Qu.:188.5   1st Qu.:2   Class :character   1st Qu.:0.0000
##  Median :410.5   Median :2   Mode  :character   Median :0.0000
##  Mean   :411.9   Mean   :2                      Mean   :0.3571
##  3rd Qu.:637.2   3rd Qu.:2                      3rd Qu.:1.0000
##  Max.   :780.0   Max.   :2                      Max.   :1.0000
##    Semester             Sex               Rubric          Rating
##  Length:196         Length:196         Length:196         1:16
##  Class :character   Class :character   Class :character   2:86
##  Mode  :character   Mode  :character   Mode  :character   3:88
##                                                           4: 6
##
##
```

```
16/196
```

```
## [1] 0.08163265
```

```
86/196
```

```
## [1] 0.4387755
```

```
88/196
```

```
## [1] 0.4489796
```

```
6/196
```

```
## [1] 0.03061224
```

```
fr3 <- fall[fall$Rater==3,]
summary(fr3)
```

```
##        X             Rater     Artifact           Repeated
##  Min.   :  1.0   Min.   :3   Length:189         Min.   :0.0000
##  1st Qu.:149.0   1st Qu.:3   Class :character   1st Qu.:0.0000
##  Median :374.0   Median :3   Mode  :character   Median :0.0000
##  Mean   :372.9   Mean   :3                      Mean   :0.3704
##  3rd Qu.:598.0   3rd Qu.:3                      3rd Qu.:1.0000
##  Max.   :741.0   Max.   :3                      Max.   :1.0000
##    Semester             Sex               Rubric          Rating
##  Length:189         Length:189         Length:189         1: 15
##  Class :character   Class :character   Class :character   2:116
##  Mode  :character   Mode  :character   Mode  :character   3: 55
##                                                           4:  3
##
##
```

```
15/189
```

```
## [1] 0.07936508
```

```
116/189
```

```
## [1] 0.6137566
```

```
55/189
```

```
## [1] 0.2910053
```

```
3/189
```

```
## [1] 0.01587302
```

Raters tend to give higher ratings in the fall semester.