

3D Visualization and Prediction of Spine Fractures

Progress Report #2

February 14, 2016

Jiayuan Shi, Jing Liu,
Qianyun Zou, Su Yang

Background

In the project “3-D Visualization and Prediction of Spine Fractures”, Professor Morgan and her group have developed a method of measuring the deformations that occur throughout a human vertebra (one of the bones in the spine) as they compress the vertebrae to failure in the laboratory. The prior studies of human vertebrae failure are using numerical simulation with Finite Elements (FE) methods, whose prediction accuracy has not been assessed until now. The goal of Professor Morgan and her group is to perform this assessment by comparing the FE simulations predictions to the measurements on an actual crushed vertebra.

Figure A and B below show the experimental result for one vertebra, and Figure 3C-J show FE predictions for the same vertebra, for eight different types of simulations. These eight simulations can be divided into two types and four methods. There are two types of tissues used in the simulations, Crushable Foam and von Mises. Under each type of tissue, there are four simulation methods applied: Experimentally Matched FE, Idealized FE, IVD-Generic FE, and IVD-Specific FE.

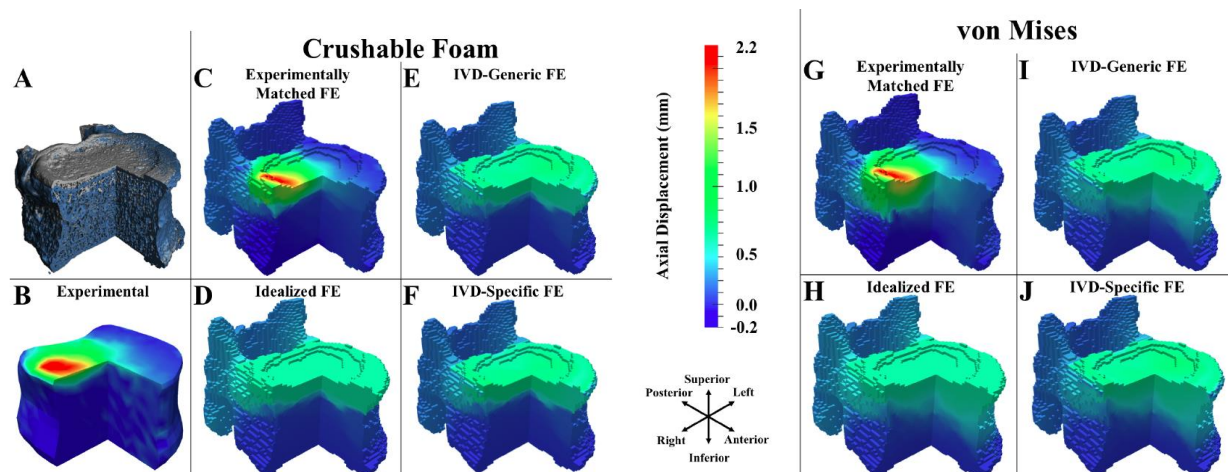


Figure A-J from Prof. Elise Morgan

Research Questions

Based on these results, Professor Morgan wants to know:

- i. How similar is each of the simulation results to the experimental results?
- ii. Are some of the eight simulations more similar than others to the experimental measurement?

Dataset

There are two types of compressions. The first is called Union, which means compression only, while the second is called Flexion, which is a combination of flexion and compression. Each compression contains data from 14 specimens, so the whole data set contains data from 28 specimens. The specimen has been divided into several grids, and one single number is provided as the deformation of one grid using region-averaging method.

Below is displayed a part of the dataset. The first line is the ID for one specimen, under which are names for simulations. There are eight simulations under each specimen (two tissue types X four models = eight distinct simulation settings). For each simulation, there are three columns: the first column is grid ID, the second column is the experimental result, and the third column is the simulation result.

	A	B	C	D	E	F	G
1	513						
2	CFisoKop_Flex			CFisoKop_Step			G
3	2	0.3973	0.8273	2	0.3973	0.5762	2
4	3	0.4058	0.8322	3	0.4058	0.5504	3
5	6	0.5779	0.7852	6	0.5779	0.6516	6
6	7	0.5399	0.7719	7	0.5399	0.6465	7
7	9	0.2495	0.5297	9	0.2495	0.5853	9
8	10	0.2818	0.4894	10	0.2818	0.5514	10
9	11	0.2776	0.6317	11	0.2776	0.6049	11
10	12	0.1264	0.6982	12	0.1264	0.6134	12
11	13	0.0022	0.4779	13	0.0022	0.5626	13
12	14	0.0627	0.4244	14	0.0627	0.523	14
13	15	0.0799	0.5188	15	0.0799	0.5412	15
14	16	0.1082	0.4091	16	0.1082	0.439	16
15	17	-0.0247	0.4511	17	-0.0247	0.5403	17

Figure of data set provided by Prof. Elise Morgan

Analysis

■ Visualization

For global comparison, we started by looking into distributions of differences between simulation and experimental results. At this stage, our analysis is ignoring the 3-D structure of the data, which means that we are only comparing the values of deformation of grids, regardless of where they are on the vertebra. Below are the plots of Flexion 513 and Union 581.

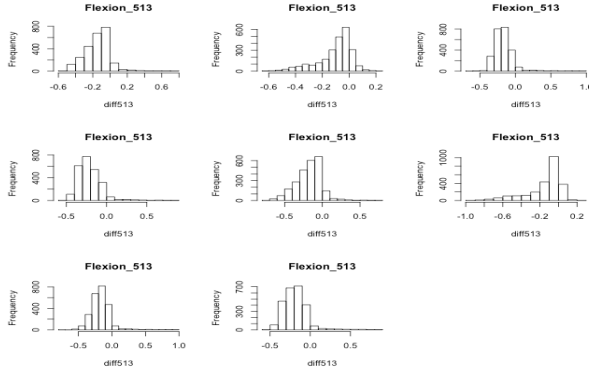


Figure K. Histograms of Vertebra Sample 513

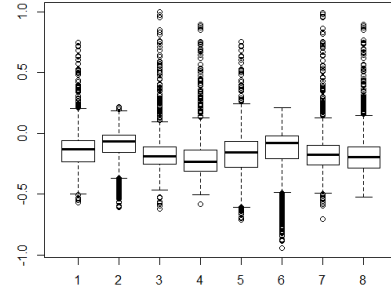


Figure L. Boxplots of Vertebra Sample 513

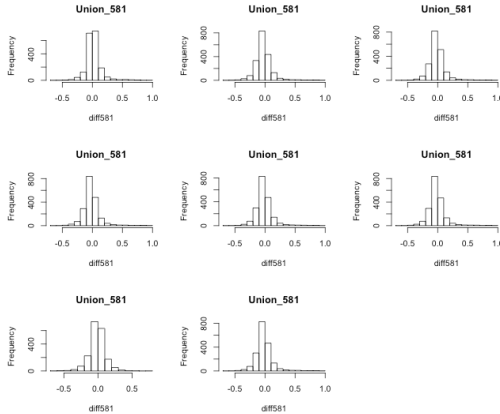


Figure M. Histograms of Vertebra Sample 581

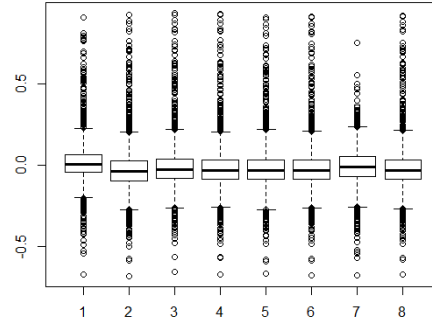


Figure N. Boxplots of Vertebra Sample 581

Results in Flexion 513 shows that the difference are not normally distributed, instead they are either left-skewed or right-skewed and there seems no patterns between. We also notice that in most cases, simulation results have higher value than experimental results. For Union 581, distributions of the difference are approximately normal. Thus, based on the above plots, we might anticipate that the performances of the eight simulations are different within the Flexion group, and similar in the Union group. To confirm this, we then move on to two-way ANOVA.

■ Two-Way ANOVA

As we mentioned before, the eight simulations differ in type and method, thus these two factors may have impact on their performances. Therefore, a two-way ANOVA is conducted to statistically test this hypothesis.

We use the mean difference as one measure of the distance between experimental and simulation results for displacements, which is,

$$\frac{1}{n} * \sum (Experimental - Simulation)$$

Another measure of distance we consider is the squared error, which is,

$$\sum (Experimental - Simulation)^2$$

In order to examine whether the types of tissue and methods of simulation have influence on mean difference and squared error between the experimental and simulation results, we conduct two-way analysis of variance (ANOVA) for both Flexion group and Union group. The two-way ANOVA not only aims at assessing the main effect of type and method but also if there is any interaction between them. The two-way ANOVA procedure, conducts a hypothesis tests to check if the factors type, method and their interactions have significant influence on the mean difference or squared error. We have a null hypothesis of non-significance, and if the resulting p-value of one factor is less than 0.05, then with 95% confidence, we can reject the null hypothesis, and conclude that factor has significant influence on the the mean difference or squared error.

Below we display the output of the two-way ANOVA procedure for Flexion group, to test the influence of type of tissue on mean difference. The procedure is based on a linear regression model which predicts the mean difference from type. The null hypothesis is that there is no significant influence of type on the mean difference between experimental and simulation results. The procedure outputs in the last column of “flexion.type” a p-value of 0.00147, which is less than 0.05, so with with 95% confidence, we can reject the above null hypothesis. Hence, type in the Flexion group has significant influence on its mean difference. Similarly, we can see in the Flexion group, not only the type, but method also has significant influence on its mean difference. However, their interaction does not influence the mean difference significantly because of its large p-value of 0.77284.

```
# Flexion: influence of types and methods on mean difference
flexion.mean.fit <- aov(flexion.mean ~ flexion.type*flexion.method)
summary(flexion.mean.fit)
#Df Sum Sq Mean Sq F value Pr(>F)
#flexion.type          1 0.0639 0.06391 10.684 0.00147 **
# flexion.method        3 0.0573 0.01912  3.196 0.02657 *
# flexion.type:flexion.method 3 0.0067 0.00223  0.373 0.77284
#Residuals            104 0.6221 0.00598
```

For each of Flexion and Union group, we conduct two two-way ANOVA models, one is for test influence on mean difference and the other is for test influence on squared error. We get the following results in the table below,

	Flexion Group			Union Group		
	Type	Method	Interaction	Type	Method	Interaction
Mean Difference	Significant	Significant	Non-significant	Non-significant	Non-significant	Non-significant
Squared Error	Significant	Non-significant	Non-significant	Non-significant	Non-significant	Non-significant

Table A. Significance Results from ANOVA

From the above results, for the three significant ones (in Flexion group, the effect of both type and method on mean difference, and the effect of type on squared error), we can also use one-way ANOVA procedure to re-confirm that they are indeed significant. The results from the one-way ANOVA are the same as those from the two-way ANOVA considering only main effects.

The results from the ANOVA procedure are not surprising given our initial exploratory data visualizations. Back to the graphs above, for Flexion group, distributions of mean differences between simulation and experimental results are not normal, instead they are either left-skewed or right-skewed away from zero (zero would indicate no difference). However, for Union group, distributions are approximately normal and centered at zero. In conclusion, via the ANOVA procedure we confirm that the levels of the factors type and method have a significant effect on the mean difference and the levels of the factor type have a significant effect on the squared error for Flexion group, but non-significant effect of type or method for Union group.

■ Tukey HSD Test

Elaborating on the ANOVA output, which shows that there is significant difference between the four methods in Flexion group on the mean difference measure of distance between simulation and experimental observations, we next want to find out where the difference lies. That is, we want to determine which methods in the group differ significantly from each other. To do this in a statistically rigorous manner, we conduct Tukey HSD Test, a post-hoc test of the ANOVA procedure.

The four levels of the factor variable method are Experimentally Matched FE, IVD-Generic FE, Idealized FE and IVD-Specific FE. The output of the Turkey HSD Test is six confidence

intervals, six since we want to consider all possible pairwise comparisons between our four levels (i.e. Experimentally Matched FE compared to IVD-Generic FE, Experimentally Matched FE compared to Idealized FE, etc.). Each confidence interval is used to perform a hypothesis test. The null hypothesis is that there is no significant difference between the two levels. If the confidence intervals does not include zero, we would reject the null hypothesis. We run this multiple testing procedure keeping the overall confidence level at 95%. The comparison between IVD-Generic FE and Experimentally Matched FE rejects the null hypothesis and reaches the conclusion that there is significant difference between them. Since the difference has a negative value, we conclude that the mean difference of Experimentally Matched FE is larger than IVD-Generic FE. Below, to the right of the confidence interval output, is a plot visualizing this conclusion. If we set the confidence level to be a little bit lower than 95%, two more groups of comparison results show up to be significant - the comparisons between Idealized FE and Experimentally Matched FE, and IVD-Specific FE and Experimentally Matched FE. The rest of the six pairwise tests are not significant. In conclusion, the mean difference of Experimentally Matched FE is always larger than the one from the other three methods while there is no significant difference between the other methods.

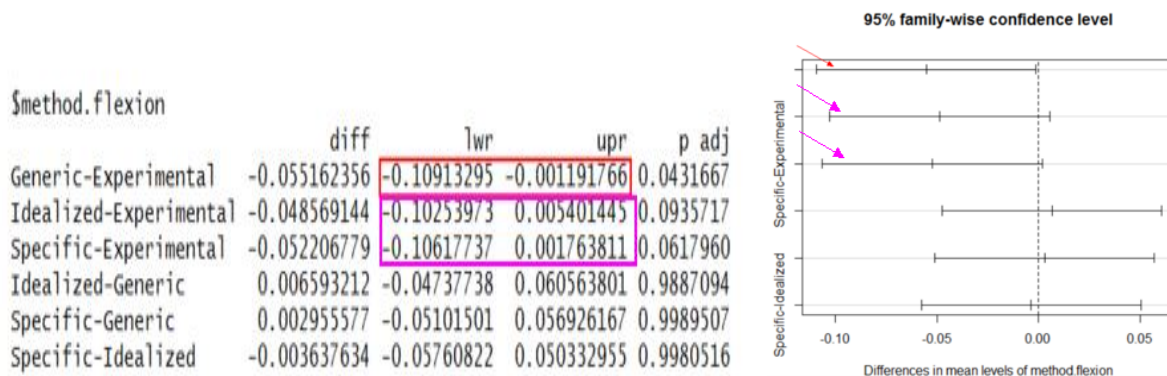


Figure O. Confidence Intervals of Tukey Results for Methods in Flexion Group

Conclusions

We have compared the simulation results to the experimental results by separating them into Flexion group and Union group. Visualization and tests show that under union compression, simulation results are very close to the experimental results and there is no statistically significant difference between either mean difference or squared error for both methods and types, which means that we cannot conclude that there is a type or a method under Union compression which is significantly better than the others. However, for Flexion group, two-way ANOVA shows that both type and method have significant influence on the mean difference. After further comparing the four methods under flexion compression using Turkey HSD Test, we conclude that since the mean difference of Experimentally Matched FE is always larger than the other methods, IVD-Generic FE, Idealized FE and IVD-Specific FE, then the latter three are

closer to the experimental results than Experimentally Matched FE simulation method. In other word, Experimentally Matched FE performs worse than the other three methods.

Possible Next Steps

Our current analysis does not take into account the 3D structure of the grid, thus spatial comparison is a natural next step. In order to do so, we can use the spatial coordinates of each grid point on the vertebra samples, but this would require a much more sophisticated model.

Appendix

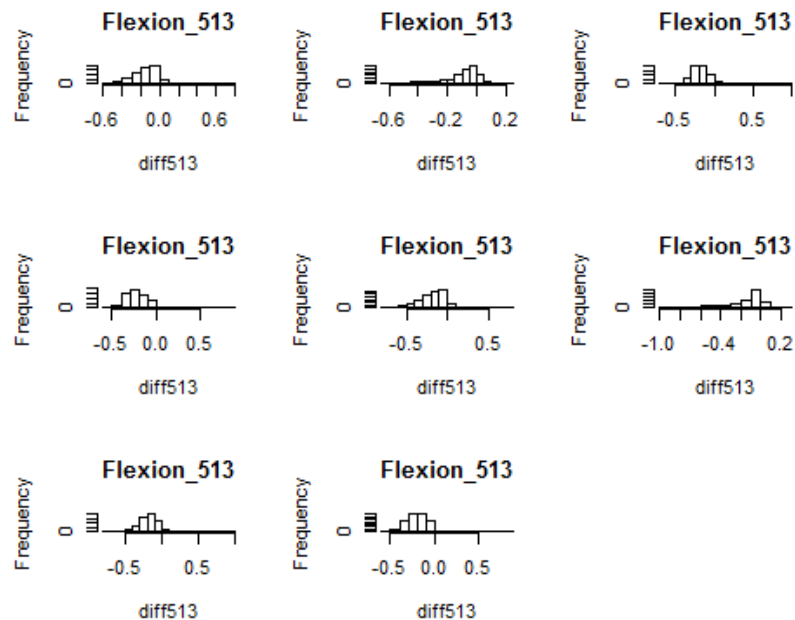
Visualization

We were using specimen 513 from Flexion group and specimen 581 from Union group as examples to study on the distribution of differences between experimental results and simulation results.

```
Flexion_513 <- read.csv("Flexion_513.csv",header=T)
Union_581 <- read.csv("Union_581.csv",header=T)
```

We created a matrix to record the differences between experimental results and eight simulation results, so it contained 8 columns and 2493 rows. We then used the data to plot histograms to see the distributions.

```
# Flexion_513
par(mfrow=c(3,3))
matrix.diff513 <- matrix(rep(0,19944), nrow=2493)
i <- 1
k <- 0
while(i < 24){
  diff513 <- Flexion_513[,i+1] - Flexion_513[,i+2]
  k <- k+1
  matrix.diff513[,k] <- diff513
  hist(diff513, main = "Flexion_513")
  i <- i+3
}
```



According to the histograms, the second and sixth simulations were highly skewed, thus we used the median difference for these two and mean difference for the rest.

```
df513 <- data.frame(matrix.diff513)
mean(df513[,1])
## [1] -0.1465736
```



```
median(df513[,2]) # the 2nd is closest to 0 one
```

```
## [1] -0.0678
```

```
mean(df513[,3])
```

```
## [1] -0.1681694
```

```
mean(df513[,4])
```

```
## [1] -0.208455
```

```
mean(df513[,5])
```

```
## [1] -0.1771904
```

```
median(df513[,6])
```

```
## [1] -0.077
```

```
mean(df513[,7])
```

```
## [1] -0.168908
```

```
mean(df513[,8])
```

```
## [1] -0.186472
```

We repeated the same loop on the Union group and plotted histogram to see their distributions.

```
# Union_581
```

```
par(mfrow=c(3,3))
```

```
matrix.diff581 <- matrix(rep(0,15512), nrow=1939)
```

```
i <- 1
```

```
k <- 0
```

```
while(i < 24){
```

```
  diff581 <- Union_581[,i+1] - Union_581[,i+2]
```

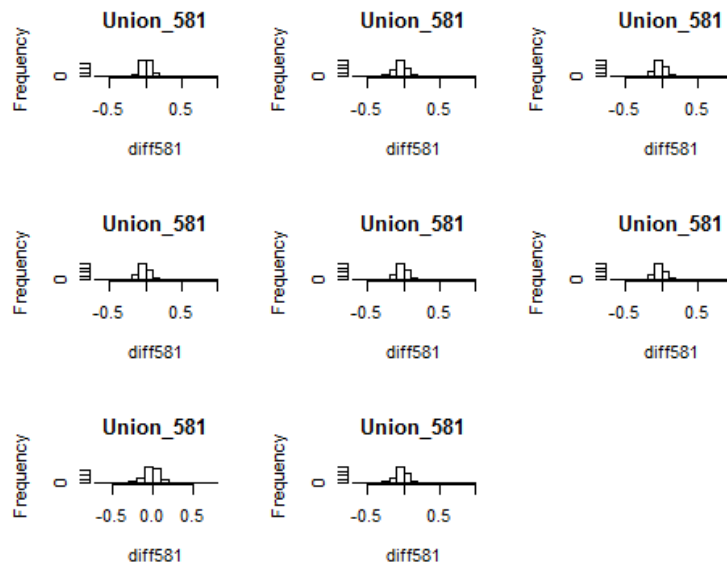
```
  k <- k+1
```

```
  matrix.diff581[k] <- diff581
```

```
  hist(diff581, main = "Union_581")
```

```
  i <- i+3
```

```
}
```



All the differences were normally distributed, thus we used mean difference for all of them.

```
df581 <- data.frame(matrix.diff581)
mean(df581[,1])
```

```
## [1] 0.01820387
```

```
mean(df581[,2])
```

```
## [1] -0.02498484
```

```
mean(df581[,3])
```

```
## [1] -0.01184528
```

```
mean(df581[,4])
```

```
## [1] -0.01739428
```

```
mean(df581[,5])
```

```
## [1] -0.01694198
```

```
mean(df581[,6])
```

```
## [1] -0.01833476
```

```
mean(df581[,7])
```

```
## [1] -0.01128592
```

```
mean(df581[,8])
```

```
## [1] -0.01891991
```

Two-Way ANOVA

We conducted two-way analysis of variance (ANOVA) for both Flexion group and Union group to examine whether the types of tissue and methods of simulation have influence on mean difference and squared error between the experimental and simulation results.

```
# read in the full data set that contains all one-dimension data points
flexion_data <- read.csv("Flexion.csv",header = T)
union_data <- read.csv("uni.csv",header = T)
# put difference of simulation and experiment results into matrix for flexion group
flexion.difference <- matrix(rep(0,279216),nrow=2493)
i <- 1
j <- 1
# use while loop to calculate the mean difference and mean squared error for flexion group
flexion.mean <- NULL
flexion.se <- NULL
while (i < 225) {
  flexion.difference[,j] <- flexion_data[i]-flexion_data[,i+1]
  flexion.mean[j] <- mean(flexion.difference[,j],na.rm = T)
  flexion.se[j] <- sum((flexion.difference[,j])^2,na.rm = T)
  i <- i+2
  j <- j+1
}
flexion.mean # mean difference for Flexion group

## [1] -0.146573566 -0.104138708 -0.168169354 -0.208455034 -0.177190413
## [6] -0.147868512 -0.168907982 -0.186471961 -0.118523414 -0.016403971
## [11] -0.170008974 -0.167416658 -0.179466478 -0.081040743 -0.181319495
## [16] -0.180374420 -0.079854393 -0.033042991 -0.094269819 -0.107792555
## [21] -0.118836284 -0.057400738 -0.124377599 -0.137192287 -0.044131127
## [26] -0.016434256 -0.058038187 -0.056425431 -0.066182992 -0.031108624
## [31] -0.071898195 -0.070503730 -0.055510710 -0.019734978 -0.097357481
## [36] -0.093479543 -0.085546450 -0.056732050 -0.098493141 -0.095939791
## [41] -0.095773486 0.009999037 -0.133301444 -0.132667188 -0.159029603
## [46] -0.037956037 -0.182380465 -0.179794986 0.027498814 -0.107361527
## [51] 0.020889273 0.019962352 -0.238825064 -0.202681949 0.013284012
## [56] 0.013036256 -0.018422834 0.014805144 -0.018473060 -0.018478745
## [61] -0.018458980 0.015600000 -0.018596255 -0.018600587 -0.050080786
## [66] 0.005467028 -0.064332611 -0.079728039 -0.103735860 -0.038618331
## [71] -0.096699599 -0.110820939 -0.026543260 -0.002174313 -0.035063649
## [76] -0.034028638 -0.045325352 -0.021103957 -0.050269484 -0.049200671
## [81] -0.117555282 -0.013443602 -0.089700361 -0.099985279 -0.145080064
## [86] -0.060756077 -0.124647413 -0.133578981 -0.101431540 -0.019300261
## [91] -0.152610287 -0.144140679 -0.278489295 -0.183200000 -0.286283342
## [96] -0.277947624 -0.054856197 -0.026956839 -0.082321460 -0.079789049
## [101] -0.085229563 -0.067926274 -0.106534978 -0.103451304 -0.188971600
## [106] -0.136961974 -0.218324549 -0.215146651 -0.322614882 -0.298329122
## [111] -0.338388528 -0.330938428

flexion.se # mean squared error for Flexion group

## [1] 108.43825 69.56601 125.58687 171.00974 153.90989 146.39255 131.61861
## [8] 146.86428 95.78464 18.28714 132.90615 130.96279 160.36935 72.81694
## [15] 146.15040 145.82172 38.07186 24.07120 47.64185 56.18610 54.73701
## [22] 37.39852 58.58412 68.75955 13.21415 8.27395 21.43628 20.81231
## [29] 19.76348 10.31805 24.94921 23.92563 41.61722 13.92168 65.23368
## [36] 60.86214 62.63556 41.79841 65.62887 62.07031 113.66871 32.24005
```

```

## [43] 135.21089 136.58673 151.02747 56.31942 170.41175 169.00875 259.16469
## [50] 185.31080 274.68154 274.42013 359.65832 378.42585 277.56040 276.97924
## [57] 16.01300 13.58482 16.17574 16.14137 16.00930 13.64339 16.18614
## [64] 16.14946 41.10382 13.38657 49.19363 61.05656 75.72020 32.67772
## [71] 62.70528 78.82927 15.84695 14.02875 17.92633 17.79986 18.50561
## [78] 14.40881 19.94705 19.66132 68.39308 18.50396 60.91276 72.77235
## [85] 107.24781 53.41309 83.09181 96.04908 68.54327 21.35631 120.73608
## [92] 113.37108 282.25177 255.52637 269.68731 254.60528 53.19779 35.88088
## [99] 71.93139 70.29777 73.04265 79.36414 87.63518 84.97729 288.88096
## [106] 234.03464 300.04574 291.22752 542.19701 642.19232 512.97650 491.96052

# put difference of simulation and experiment results into matrix for union group
union.difference <- matrix(rep(0,320096),nrow=2858)
i <- 1
j <- 1
# use while loop to calculate the mean difference and mean squared error for union group
union.mean <- NULL
union.se <- NULL
while (i < 225) {
  union.difference[,j] <- union_data[,i]-union_data[,i+1]
  union.mean[j] <- mean(union.difference[,j],na.rm = T)
  union.se[j] <- sum((union.difference[,j])^2,na.rm = T)
  i <- i+2
  j <- j+1
}
union.mean # mean difference for Union group

## [1] 1.820387e-02 -2.498484e-02 -1.184528e-02 -1.739428e-02 -1.694198e-02
## [6] -1.833476e-02 -1.128592e-02 -1.891991e-02 9.543081e-03 -9.365921e-02
## [11] -7.674132e-02 -7.222651e-02 -1.686603e-01 -1.652553e-01 -7.785624e-02
## [16] -1.844397e-01 -2.290037e-05 -9.523579e-02 -7.693949e-02 -8.044105e-02
## [21] -9.168289e-02 -9.441408e-02 -2.537533e-02 -9.876275e-02 -1.050185e-02
## [26] -1.335314e-01 -9.762740e-02 -1.043004e-01 -1.337649e-01 -1.398198e-01
## [31] -5.279142e-02 -1.524444e-01 4.691471e-02 1.386600e-01 1.285016e-01
## [36] 1.263566e-01 1.228149e-01 1.223633e-01 7.631800e-02 1.243254e-01
## [41] 2.506402e-02 1.567684e-01 1.047324e-01 1.201016e-01 1.283957e-01
## [46] 1.354144e-01 8.269771e-02 1.462168e-01 -1.270375e-02 -1.405289e-01
## [51] -1.032625e-01 -1.108355e-01 -1.447512e-01 -1.502409e-01 -7.885810e-02
## [56] -1.658458e-01 1.170256e-02 -2.284291e-01 -1.443037e-01 -1.396014e-01
## [61] -1.757283e-01 -1.733341e-01 -6.012150e-02 -1.965193e-01 5.396759e-02
## [66] 3.085589e-01 2.561301e-01 2.462396e-01 2.877649e-01 2.827113e-01
## [71] 1.977569e-01 3.196125e-01 -1.136768e-02 -1.282465e-01 -1.141813e-01
## [76] -1.138910e-01 -1.322437e-01 -1.313254e-01 -6.085300e-02 -1.358201e-01
## [81] -5.468621e-02 -4.511957e-01 -3.936149e-01 -3.936460e-01 -3.577057e-01
## [86] -3.544575e-01 -1.626323e-01 -3.871902e-01 -8.959723e-02 -1.862631e-01
## [91] -1.761372e-01 -1.751574e-01 -2.486635e-01 -2.471287e-01 -1.635912e-01
## [96] -2.507047e-01 6.654982e-02 3.523938e-01 2.915648e-01 3.084099e-01
## [101] 3.619355e-01 3.765020e-01 2.361054e-01 3.888654e-01 1.866945e-02
## [106] -1.228588e-01 -8.500899e-02 -8.358214e-02 -9.719430e-02 -9.690438e-02
## [111] -1.692846e-02 -1.011225e-01

union.se # mean squared error for Union group

## [1] 34.682506 42.822910 41.023274 41.482149 40.652777 40.984330
## [7] 27.491128 41.950922 37.628122 133.780540 122.146804 115.550235
## [13] 194.450995 186.284622 128.632267 210.862493 8.683980 38.901566
## [19] 32.481108 32.843102 38.429902 38.592316 14.113974 40.632516

```

```
## [25] 35.652128 117.981071 86.899405 93.422306 110.743329 118.548034
## [31] 56.758305 132.221725 29.055917 70.535133 64.155332 62.508161
## [37] 51.764687 51.345928 44.264041 51.807247 36.172016 146.424342
## [43] 102.830434 111.982827 105.185200 108.485985 63.538099 121.424008
## [49] 19.130106 74.761851 61.060404 60.032498 79.991378 78.466873
## [55] 34.875369 86.993325 29.119735 179.222760 104.261231 100.018769
## [61] 131.810394 128.402870 65.609223 151.078440 48.389967 373.291935
## [67] 293.240669 279.278182 234.687586 228.256455 310.002364 277.452177
## [73] 9.747391 74.695723 65.121086 66.109481 75.026522 75.409769
## [79] 46.364921 78.884079 109.999450 986.665245 813.282225 833.891216
## [85] 714.770802 723.095060 321.114590 773.849438 60.877700 164.253918
## [91] 148.783409 152.649667 212.913080 215.539335 147.904996 216.641292
## [97] 80.332582 565.994535 424.604911 465.913070 485.017739 520.023488
## [103] 379.826193 543.765267 32.845343 79.303765 53.909259 53.769655
## [109] 60.526424 61.122220 16.869765 65.135292
```

For Flexion group, we firstly constructed two matrices. They are, flexion.data.mean with three columns: type, method and mean difference;
flexion.data.se with three columns: type, method and squared error.

```
flexion.type <- c(rep(c(rep("CF",4),rep("vM",4)), 14))
flexion.method <- c(rep(c("Idealized","Experimental","Specific","Generic"),28))
# Flexion matrices
flexion.data.mean <- cbind(flexion.type, flexion.method, flexion.mean )
flexion.data.se <- cbind(flexion.type, flexion.method, flexion.se)
```

Next, we used two-way ANOVA procedure to conduct a hypothesis tests to check if the Flexion type, method and their interactions have significant influence on the mean difference or squared error.

```
# Influence of Flexion types and methods on mean difference
flexion.mean.fit <- aov(flexion.mean ~ flexion.type*flexion.method)
summary(flexion.mean.fit)

##              Df Sum Sq Mean Sq F value Pr(>F)
## flexion.type      1 0.0639 0.06391  10.684 0.00147 **
## flexion.method     3 0.0573 0.01912   3.196 0.02657 *
## flexion.type:flexion.method 3 0.0067 0.00223   0.373 0.77284
## Residuals       104 0.6221 0.00598
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The p-value of flexion.type is 0.00147, which indicates a significant influence of Flexion type on the mean difference. Similarly, the p-value of flexion.method is 0.02657, which indicates a significant influence of Flexion method on the mean difference.

```
# Influence of Flexion types and methods on squared error
flexion.se.fit <- aov(flexion.se ~ flexion.type*flexion.method)
summary(flexion.se.fit)

##              Df Sum Sq Mean Sq F value Pr(>F)
## flexion.type      1 75975 75975  5.108 0.0259 *
## flexion.method     3 18659 6220  0.418 0.7403
## flexion.type:flexion.method 3 11397 3799  0.255 0.8573
## Residuals       104 1546770 14873
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The p-value of flexion.type is 0.0259, which indicates a significant influence of Flexion type on the squared error.

For Union group, we also firstly constructed two matrices. They are,
union.data.mean with three columns: type, method and mean difference;
union.data.se with three columns: type, method and squared error.

```
type.union <- c(rep(c(rep("CF",4),rep("vM",4)), 14))
union.method <- c(rep(c("Idealized", "Experimental", "Specific", "Generic"),28))
# Union matrices
union.data.mean <- cbind(type.union, union.method, union.mean )
union.data.se <- cbind(type.union, union.method, union.se )
```

Next, we used two-way ANOVA procedure to conduct a hypothesis tests to check if the Union type, method and their interactions have significant influence on the mean difference or squared error.

```
# Influence of Union types and methods on mean difference
union.mean.fit <- aov(union.mean ~ type.union*union.method)
summary(union.mean.fit)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## type.union      1  0.003  0.00306   0.094  0.760
## union.method     3  0.015  0.00505   0.156  0.926
## type.union:union.method  3  0.023  0.00779   0.240  0.868
## Residuals      104  3.378  0.03248
```

All the p-values are larger than 0.05, which indicates non-significant influence of any Union factors on the mean difference.

```
# Influence of Union types and methods on squared error
union.se.fit <- aov(union.se ~ type.union*union.method)
summary(union.se.fit)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## type.union      1  9964   9964   0.263  0.609
## union.method     3 141505  47168   1.243  0.298
## type.union:union.method  3 159942  53314   1.405  0.245
## Residuals      104 3945671  37939
```

All the p-values are larger than 0.05, which indicates non-significant influence of any Union factors on the mean difference.

Tukey HSD Test

We next conducted Tukey HSD Test to determine which methods in the group differ significantly from each other.

```
# Tukey test on mean difference for methods at 95% confidence level
flexion.mean.Tukey <- TukeyHSD(flexion.mean.fit, conf.level=0.95)
flexion.mean.Tukey

## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = flexion.mean ~ flexion.type * flexion.method)
##
```

```
## $flexion.method
##           diff      lwr      upr    p adj
## Generic-Experimental -0.055162356 -0.10913295 -0.001191766 0.0431667
## Idealized-Experimental -0.048569144 -0.10253973  0.005401445 0.0935717
## Specific-Experimental -0.052206779 -0.10617737  0.001763811 0.0617960
## Idealized-Generic      0.006593212 -0.04737738  0.060563801 0.9887094
## Specific-Generic       0.002955577 -0.05101501  0.056926167 0.9989507
## Specific-Idealized     -0.003637634 -0.05760822  0.050332955 0.9980516
##

# plot to make it clear
plot(flexion.mean.Tukey)
```

