Martha Foye, Dave Hiltbrand, and Allen Phelps

Dr. Cuixian Chen

STT 592

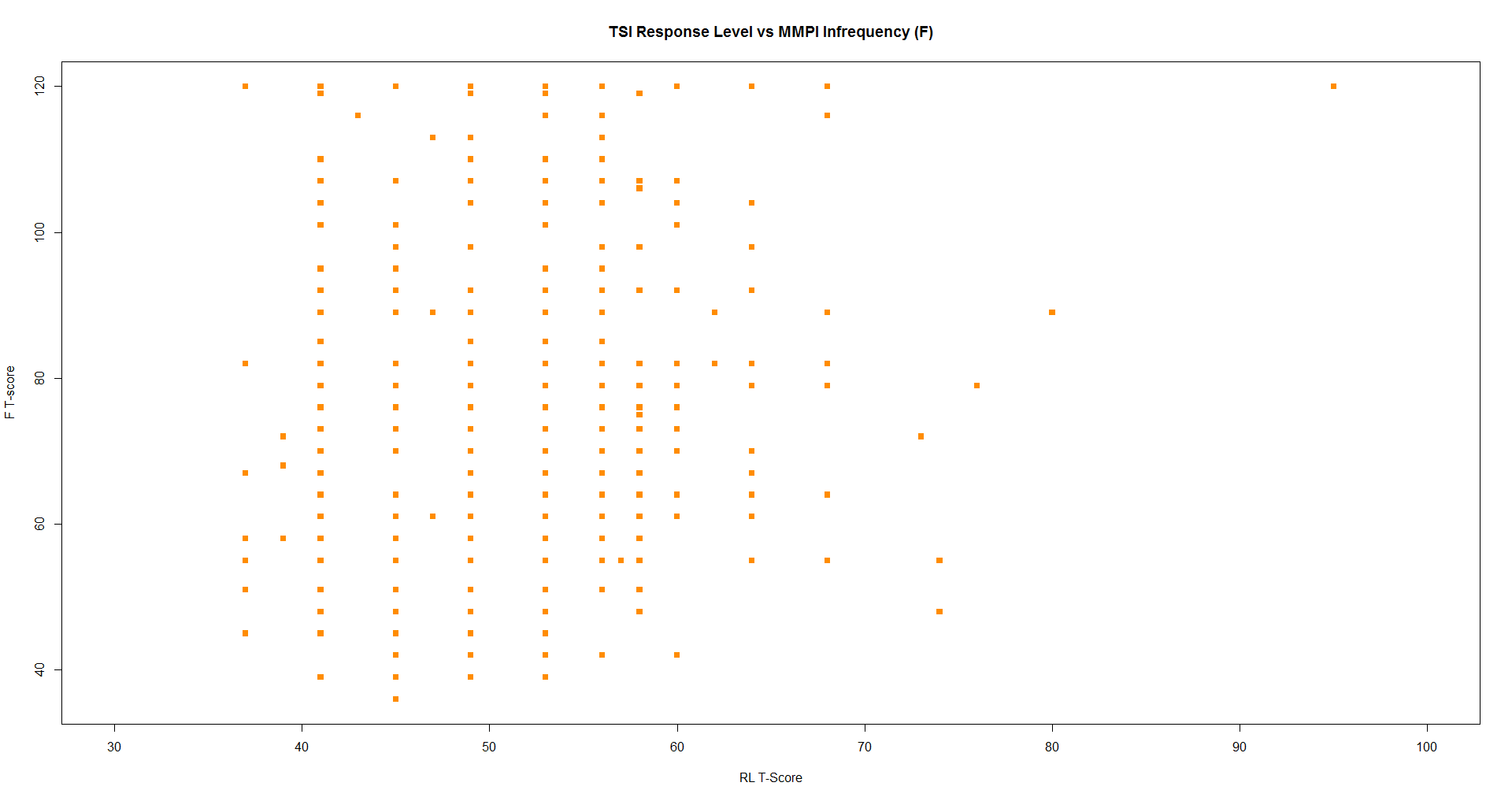
15 September 2016

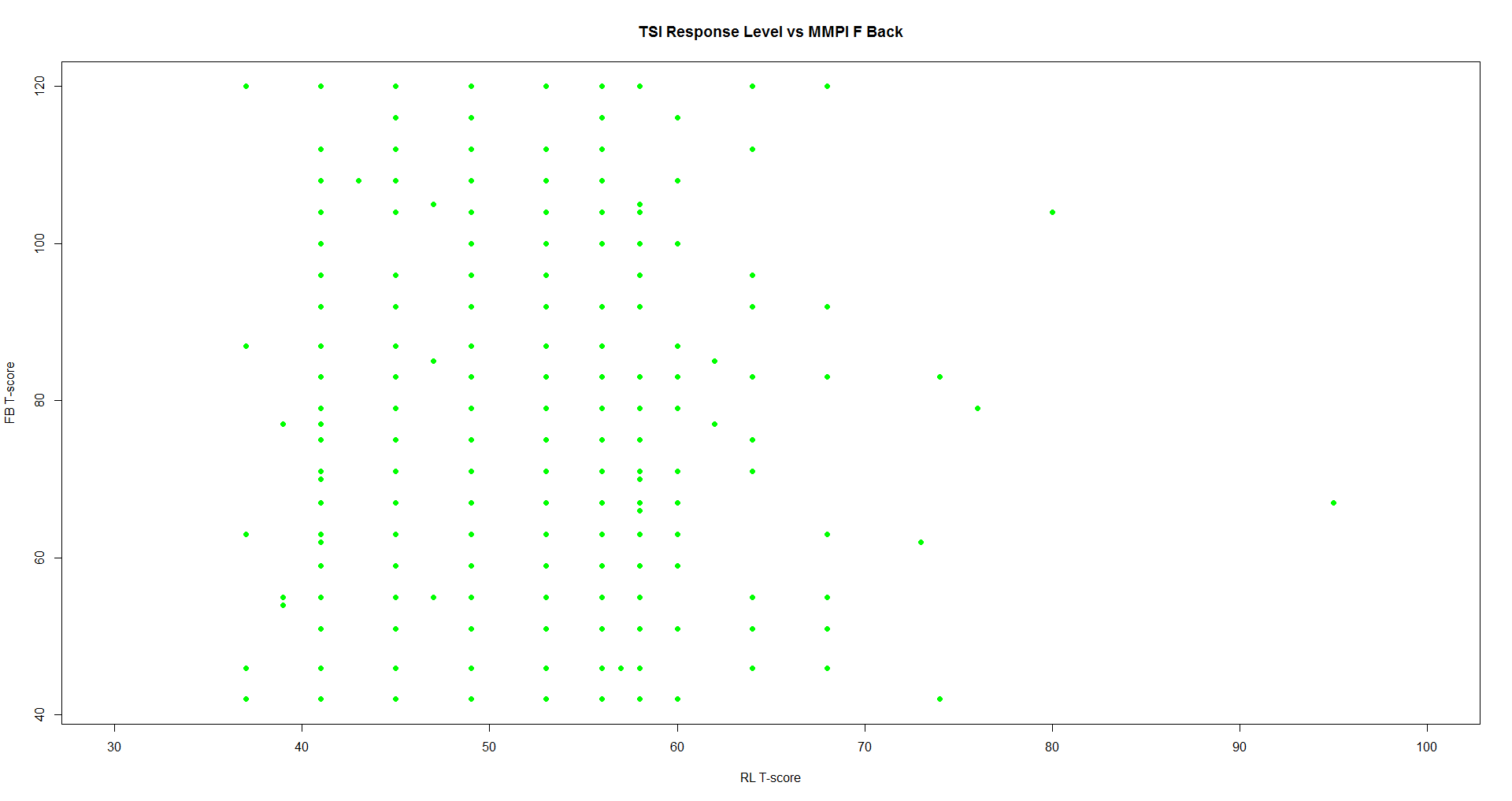
**Section 1: Linear Regression**

To evaluate the reliability of our dataset, we began with the reasonable assumption that tests or scales of tests designed to measure the same thing would agree by measure of positive, perhaps linear, correlation. For example, we assumed that a high score on the Beck Anxiety Inventory would correlate well with a heightened anxiety level on the Trauma Symptom Inventory (TSI) Anxious Arousal scale. Then, we would have a sense of consistency in participant response to counter suspicions of clinical invalidity due to self-report. A similar evaluation of validity itself follows from the comparison between tests of scales designed to measure inconsistency, malingering, and superlative self-presentation. Finally, in particular, we compare the clinical scales of both the TSI and Minnesota Multiphasic Personality Inventory to the Test of Memory Malingering to test whether or not it is indeed independent of the mental pathology based on our psychometrics, which would confirm its designed utility in validating malingering.

**a. Pairwise Comparison (Validity between Tests)**

Both the Trauma Symptom Inventory (TSI) and the Minnesota Multiphasic Personality Inventory (MMPI) have built-in validity scales. Our hypothesis states that responses to the various types—inconsistent, “faking good”, and “faking bad”—would match up between the two tests, most likely indicating a linear relationship. After using scatterplots to look for relationships between the variables, it was easy to determine that no such relationship exists. In fact, it appears that no relationship exists at all. Further investigation is required to understand why if the two tests are measuring the same type of validity that a participant would not score similarly on both. Take the TSI Response Level versus the MMPI Infrequency and F Back and scales for instance. All are designed to measure malingering, yet we observe the full range of values at each t-score of the other:

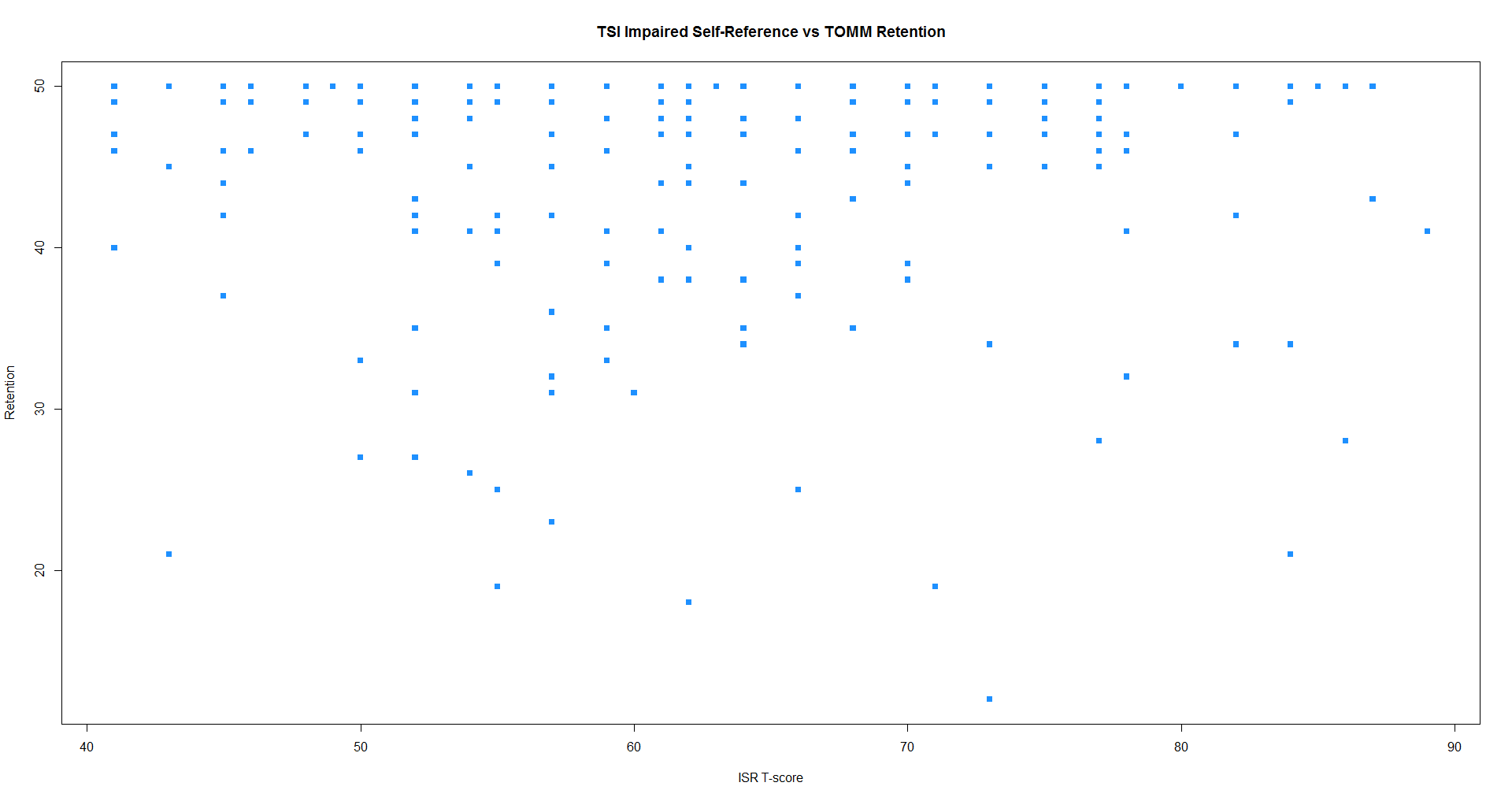


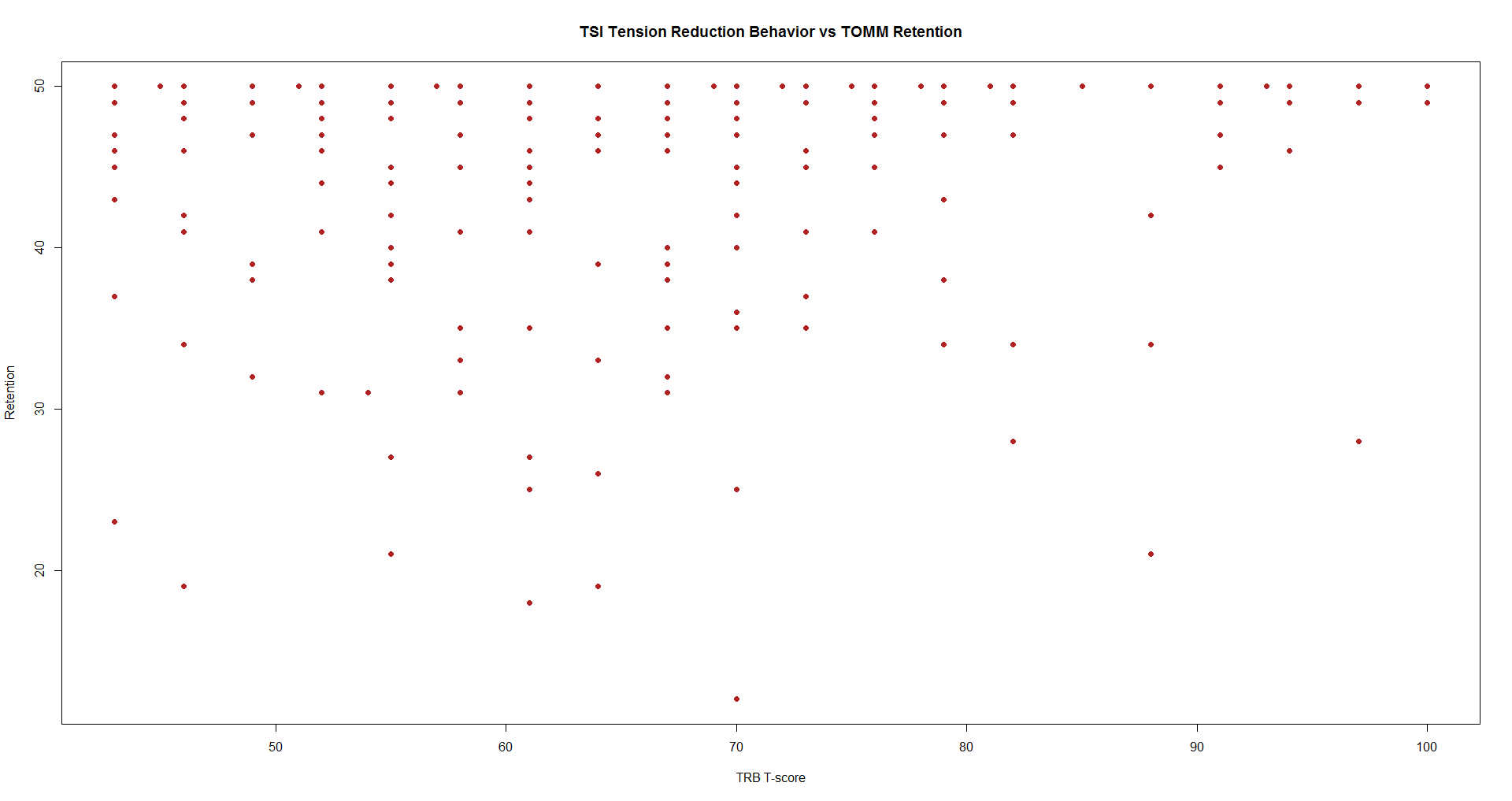


Combined with linear coefficients of correlation, 0.1048411 and 0.06890269, respectively, did not provoke us to pursue modeling one after the other, much less linearly. Unfortunately, we found the same to be true of the TSI Inconsistent Response versus the MMPI Variable Response Inconsistency as well as the TSI Atypical Response vs the MMPI Lie, Defensiveness, and Superlative Self-Representation scales. To say the least, it was disappointing not to find consistency between tests among the very measures of similar validity.

**b. Pairwise Comparison (Validity of TOMM)**

Contrarily, the lack of dependency between the clinical scales of both the Trauma Symptom Inventory and the Minnesota Multiphasic Personality Inventory and the Test of Memory Malingering (TOMM) was refreshing to observe. The only we pattern observed, as expected, was a clustering of retention between 45 and 50 over the full range of t-scores for each clinical scale. Our expectation was based on the prior knowledge that the TOMM was designed to elicit this range of responses regardless of even severe mental pathology. Indeed, less than 13% of the participants retained less than 45. Since the TSI and MMPI each have ten clinical scales, we offer only a couple of scatterplots to illustrate the clusters:

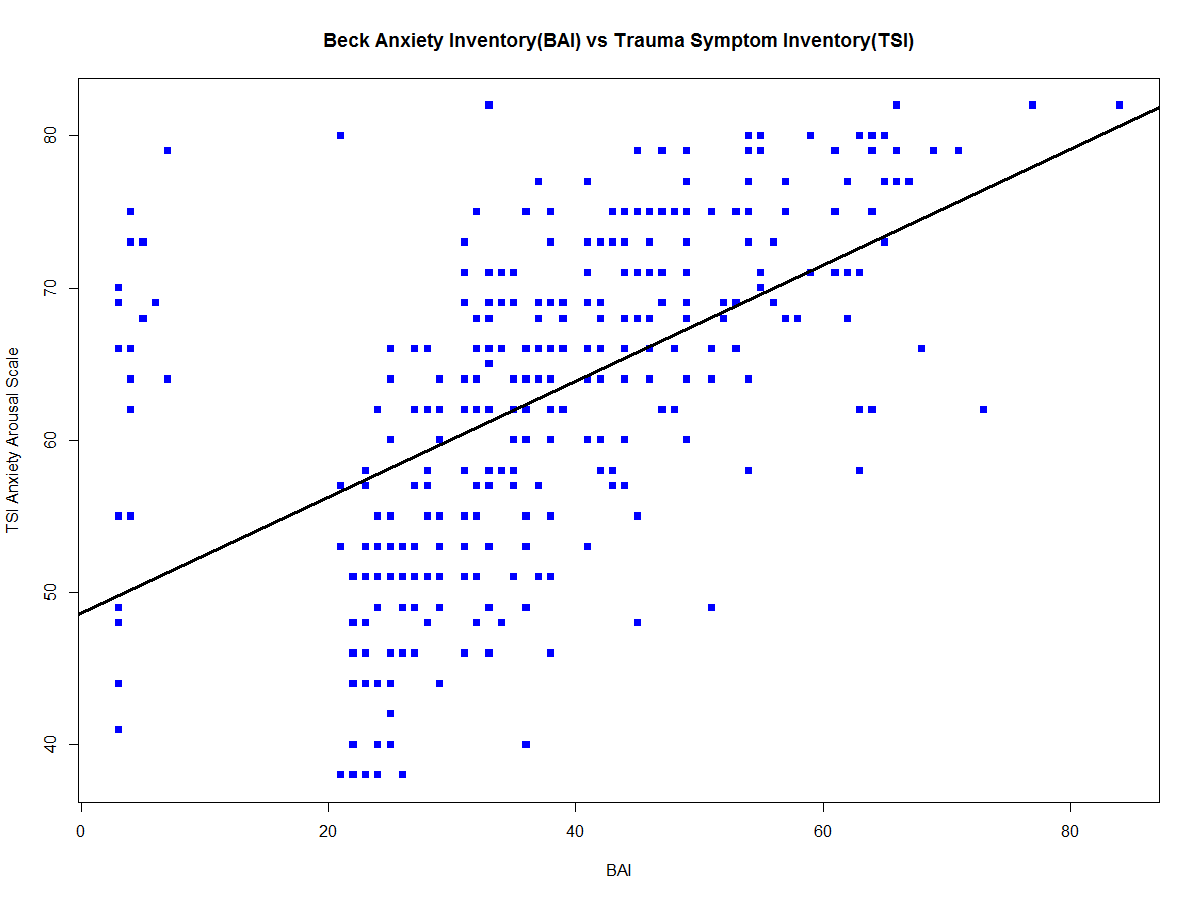




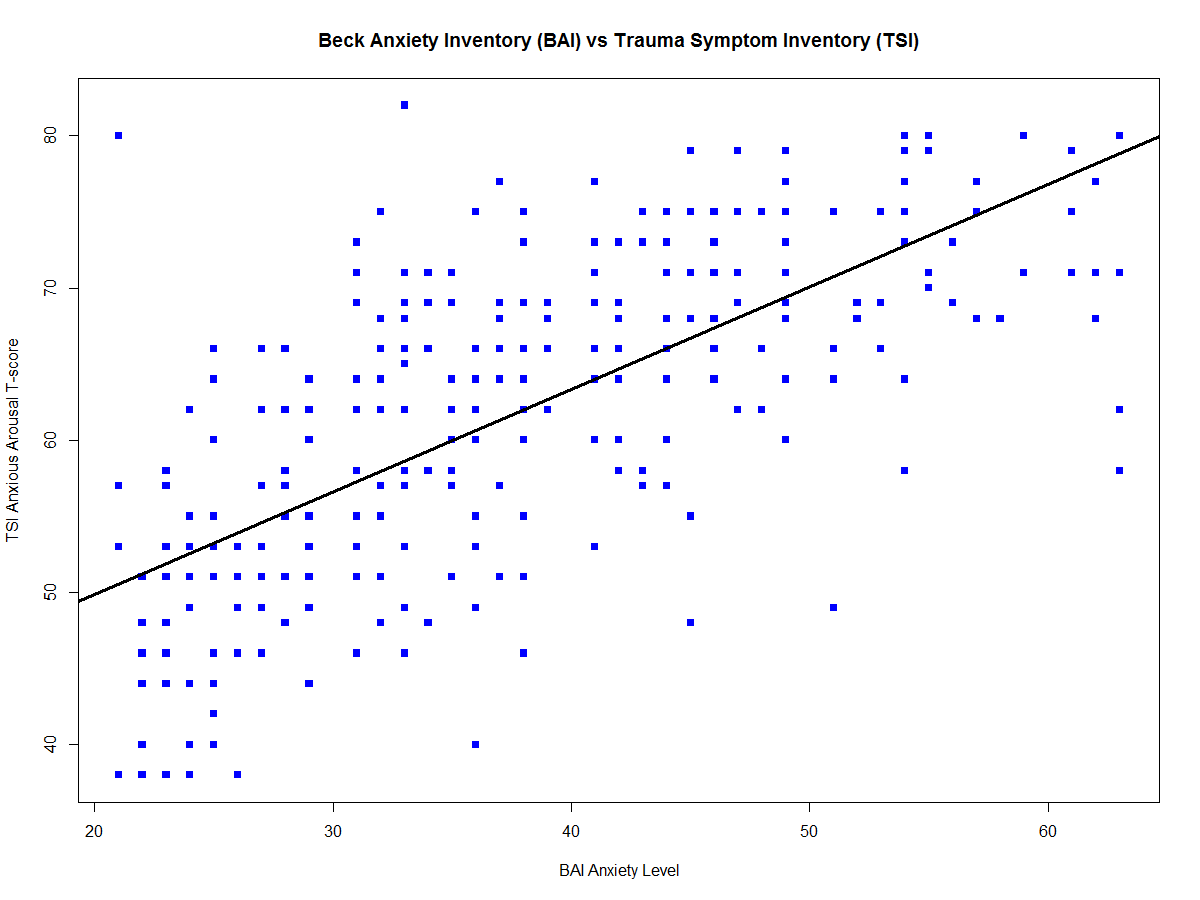
Again, we do not consider modeling, but no other pattern is desired in this case. TOMM does not seem to discriminate between variables, so it gives us at least one sense of consistency that we were looking for in terms of reliability within dataset.

**c. Simple Linear Regression (Clinical between Tests)**

It was not until we compared similar clinical measures of psychopathology that we began to observe linear patterns between similar test variables. Simple linear models of anxiety and depression were built between the Beck Anxiety Inventory (BAI), the Zung Depression Scale, the Anxious Arousal (AA) and Depression scales of the Trauma Symptom Inventory (TSI), and the corresponding Psychasthenia and Depression scales of the Minnesota Multiphasic Personality Inventory. To illustrate the modeling process, consider the TSI AA t-scores versus the BAI anxiety levels. As we hypothesized, a scatterplot revealed a positive linear pattern but curiously only for BAI scores beyond about 20:



Indeed, a summary of the linear model fit including all of the data points reported only a 0.5346 coefficient of correlation between the two measures of anxiety, yet the relationship appears much stronger past this threshold. Further inquiry revealed that a Beck Anxiety Inventory (BAI) score of 19 separates intervals of mild to moderate from moderate to severe anxiety. Moreover, 63 is the maximum possible score, which discouragingly invalidated 5.13308% of the BAI data entry; thus, to zero in on the central pattern, we fit a linear model to the valid, moderate and severe entries vs the Trauma Symptom Inventory (TSI) Anxious Arousal (AA) t-scores:





For every unit increase in BAI score (), the TSI AA t-score () increases by 0.67381 on average. The additional research aided toward improving the final model, the summary of which now reported a much more moderate 0.6867 coefficient of correlation that we expected to see. That is if the participants responded consistently to both inventories.

The remaining clinical comparisons proceeded in much the same manner with similar results of moderate correlation among the abovementioned measures of anxiety and depression, respectively:

bai2 aa pt

bai2 1.0000000 0.6878656 0.5984299

aa 0.6878656 1.0000000 0.6466019

pt 0.5984299 0.6466019 1.0000000

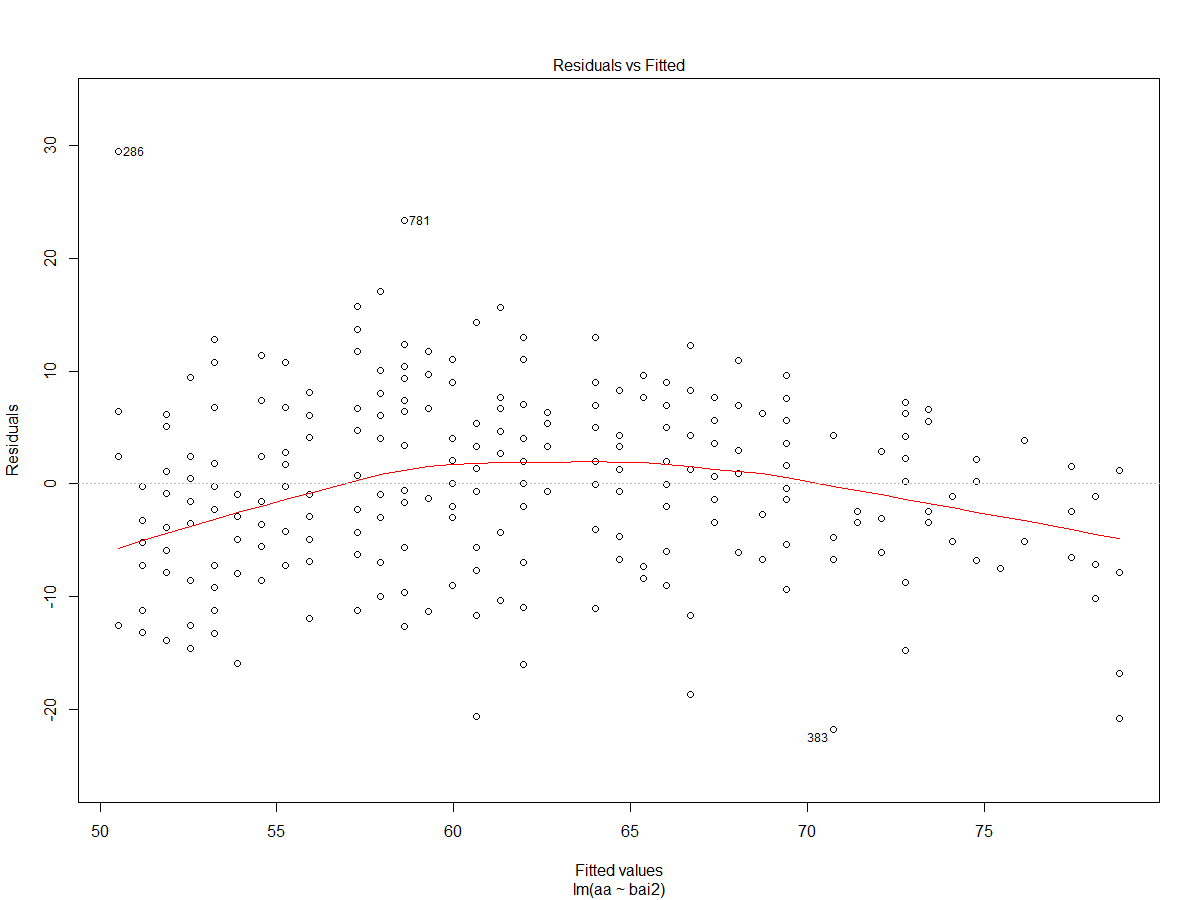
zds2 d dm

zds2 1.0000000 0.6944469 0.7539836

d 0.6944469 1.0000000 0.6226958

dm 0.7539836 0.6226958 1.0000000

Each comparison underwent model adequacy checking, especially the assumption of normally and independently distributed residuals with constant variance. In addition to visual analysis of the fitted values versus residual plots, a formal Anderson-Darling test for normality was performed—for good measure—and satisfied in each case. Take the case presented in this section for example:



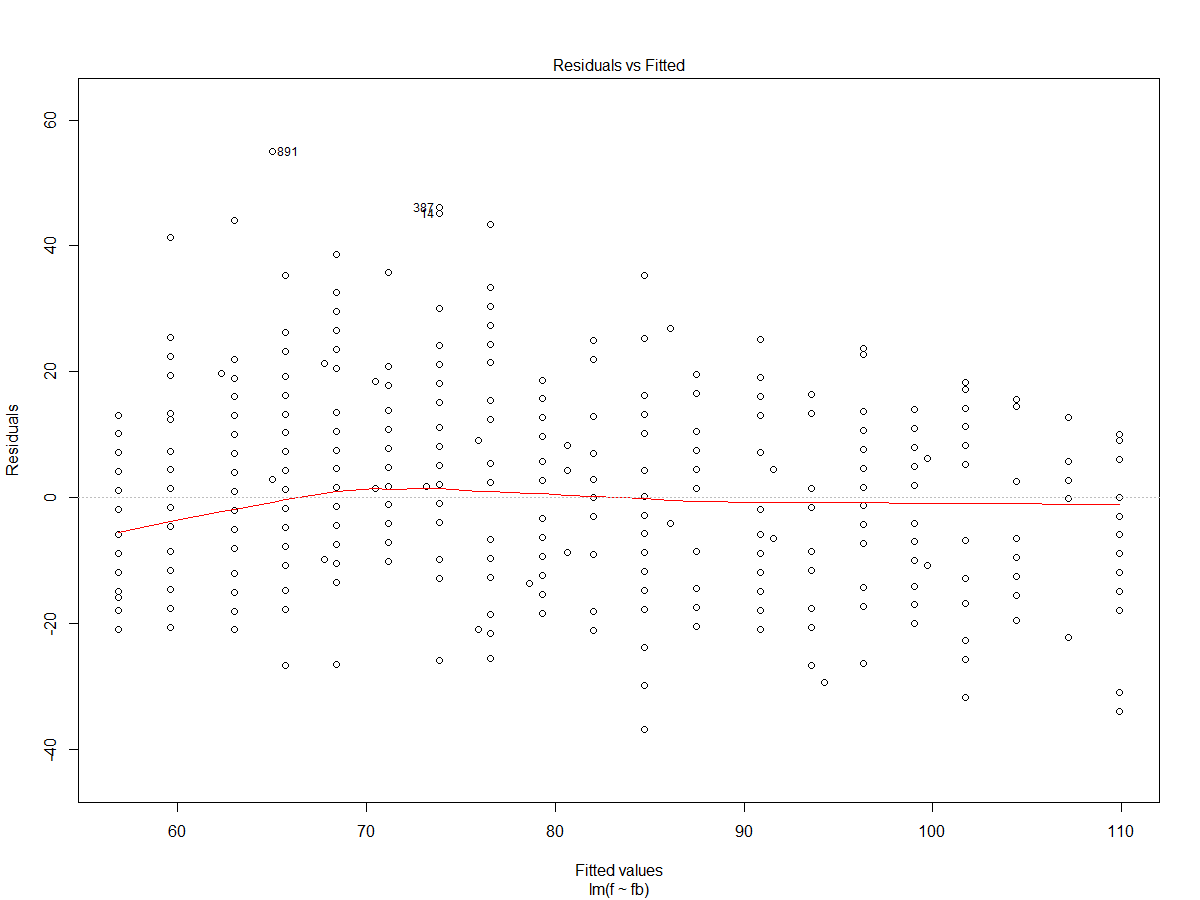
Anderson-Darling normality test

data: y2$residuals

A = 0.63278, p-value = 0.09829

**d. Multiple Linear Regression (Validity within MMPI)**

Whereas the 100-item Trauma Symptom Inventory has one measure of inconsistency, malingering, and superlative self-representation, the much lengthier 567-item Minnesota Multiphasic Personality Inventory (MMPI) has three scales for each of these validity measures; therefore, another data consistency of interest is the agreement among similar validations within the MMPI itself. Among the more interesting comparisons was between the Infrequency (F) and F Back (FB) t-scores, which measure malingering, or “faking bad”, in the first and second half of test, respectively. While an initial scatterplot shows somewhat of a positive linear pattern supported by a moderate 0.7672 summary coefficient of correlation when fit with regression, the residuals were difficult to normalize during model adequacy checking:



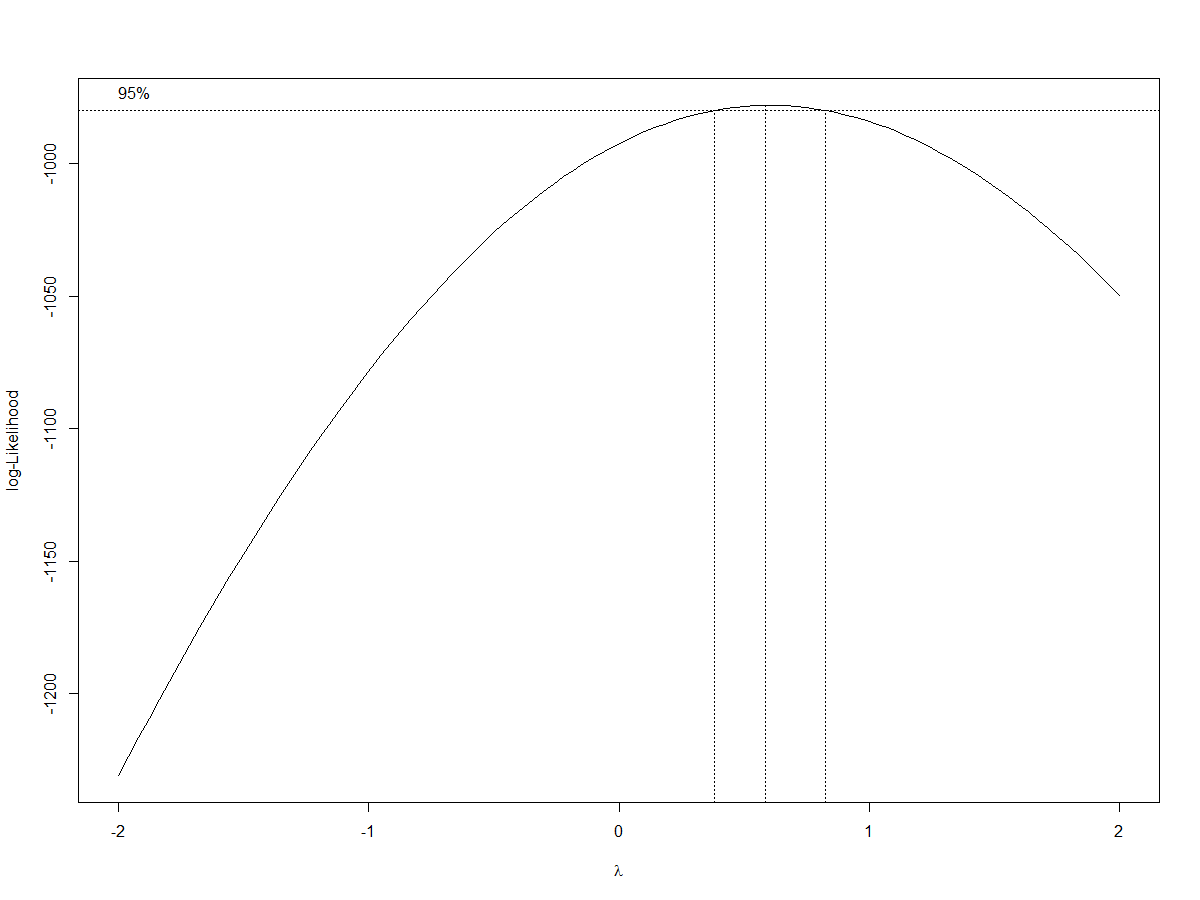
Anderson-Darling normality test

data: y17$residuals

A = 2.0071, p-value = 4.095e-05

First, we attempted to remove the possible influence of the six residuals outlying the range of 40 about 0 to no avail. Several more attempts to narrow this range resulted in the removal of too many data points, none of which resulted in subsets of residuals that failed to reject the Anderson-Darling test of normality. Next, we tried to model the Infrequency t-scores as the F Back t-scores, their squares as an indicator function, and the interaction of the two, noting a possible quadratic pattern in the residual plot. While this approach did result in a better model, each parameter of which was significant, the violation of normality persisted. Finally, we turned to transformation in effort to exhaust the exploration of suspect signal in residual curvature.

For this, we used the boxcox function of the standard MASS library in R to identify—with 95% confidence—the best power of the Infrequency (F) t-scores that could be linearly modeled by the original F Back (FB) t-scores:





Based on the resultant plot, the square root of the F t-scores () was modeled by the FB t-scores () just as moderately as before () but finally with normalized residuals. Although the final model does not include multiple predictor variables, as the section title suggests, the modeling process did consider multiple linear regression until the residuals forced the consideration of transformation, which forfeits the ideality of direct interpretation. According to our model, for every unit increase in FB t-score, the square root of the F t-score increases by 0.038578 on average. The main reason for choosing this model was to elaborate on the challenge it presented in terms of residual analysis.

**Section 2: Logistic Regression**

Arguably the most important categorical variable in the dataset is the clinical diagnosis of participants with Posttraumatic Stress Disorder (PTSD or not) by their attending neuropsychologist. Since the Trauma Symptom Inventory (TSI) was designed to assess PTSD and other traumatic sequelae, we suppose there exists a logit model comprised of a combination of its scales as predictor variables such that—given the corresponding t-scores—the probability that a participant has been diagnosed is modeled well. Then, an accurate fit of the self-reported responses that agrees with professional diagnostics would support the reliability of the data in our evaluation. Unlike the simple and multiple linear regression in the first section, the logistic regression below considers 13 predictors, one per scale of the TSI; thus, the leaps package in R was helpful in determining which scales, or any interaction between, would be significant in modeling the desired probability.

Forward selection, backward elimination, and both directions were searched with the step function of the leaps package to determine that the most parsimonious logit model included the t-scores for Anxious Arousal (AA), Defensive Avoidance (DA), and the interaction of the two Trauma Symptom Inventory (TSI) scales:

,

where  is the probability that a participant has been diagnosed with Posttraumatic Stress Disorder given his or her AA t-score () and DA t-score () on the TSI. Although backward elimination determined a different model with higher overall, sensitivity, and specificity accuracy rates, it was regrettably over parameterized and, thus, riddled with multicollinearity. Meanwhile, the simpler model resulted in the following confusion matrix:

ptsd

pred FALSE TRUE

FALSE 142 47

TRUE 133 293

From the table, it follows that the overall accuracy, true-negative, and true-positive rates are 70.73171, 51.63636, and 86.17647 percent, respectively, modeling the dataset as a whole.

In order to test the moderate ability of the Trauma Symptom Inventory Anxious Arousal and Defensive Avoidance t-scores appeared to have to predict the probability that a participant has been diagnosed with PTSD, a training model based on eighty percent of the dataset was fit to test the remaining twenty percent with the following results:

ptest

pred FALSE TRUE

FALSE 27 3

TRUE 32 61

The overall and sensitivity accuracy rates improved to 71.54472 and 95.31250 percent, respectively, whereas specificity decreased to 45.76271%; therefore, the model is not successful enough to obviate professional diagnosis since it incorrectly predicts that over half of the participants without posttraumatic stress belong among the participants who have been diagnosed with Posttraumatic Stress Disorder (PTSD). In terms of compensation that veterans receive to treat the mental aftermath of deployment, misdiagnosis of this magnitude could be quite costly. Nonetheless, the level of overall accuracy observed supports the supposition that Trauma Symptom Inventory is assessing PTSD as intended, which further supports the reliability of the data on which the prior was based.

**Appendix: R Code**

dat=read.csv("stt592dat.csv", header=T)

attach(dat)

install.packages('nortest')

library(nortest)

library(MASS)

#Martha Simple Linear Regression (Clinical b/t Tests)

plot(bai,aa,col='blue',main='Beck Anxiety Inventory(BAI) vs Trauma Symptom Inventory(TSI)',pch=15,xlab='BAI',ylab='TSI Anxiety Arousal Scale')

y=lm(aa~bai)

abline(y,col='black',lwd=3)

summary(y)

par(mfrow=c(2,2))

plot(y);ad.test(y$residuals)

dev.off()

bai2=ifelse(bai>18&bai<64,bai,NA)

plot(bai2,aa,col='blue',main='Beck Anxiety Inventory (BAI) vs Trauma Symptom Inventory (TSI)',pch=15,xlab='BAI Anxiety Level',ylab='TSI Anxious Arousal T-score')

y2=lm(aa~bai2)

abline(y2,col='black',lwd=3)

summary(y2)

par(mfrow=c(2,2))

plot(y2);ad.test(y2$residuals)

dev.off()

bai3=ifelse(bai>63,bai,NA)

(length(bai)-sum(is.na(bai3)))/(sum(!is.na(bai)))

plot(bai2,pt,col='blue',main='Beck Anxiety Inventory (BAI) vs Minnesota Multiphasic Personality Inventory (MMPI)',pch=15,xlab='BAI Anxiety Level',ylab='MMPI Psychathenia T-score')

y3=lm(pt~bai2)

abline(y3,col='black',lwd=3)

summary(y3)

par(mfrow=c(2,2))

plot(y3);ad.test(y3$residuals)

dev.off()

plot(zds,d,col='green',pch=15,xlab='ZDS Level of Depression',ylab='TSI Depression T-score',main='Zung Depression Scale (ZDS) vs Trauma Symptom Inventory (TSI)')

y4=lm(d~zds)

abline(y4,col='black',lwd=3)

summary(y4)

par(mfrow=c(2,2))

plot(y4);ad.test(y4$residuals)

dev.off()

zds2=ifelse(zds>19,zds,NA)

plot(zds2,d,col='green',main='Zung Depression Scale (ZDS) vs Trauma Symptom Inventory (TSI)',pch=15,xlab='ZDS Level of Depression',ylab='TSI Depression T-score')

y5=lm(d~zds2)

abline(y5,col='black',lwd=3)

summary(y5)

par(mfrow=c(2,2))

plot(y5);ad.test(y5$residuals)

dev.off()

zds3=ifelse(zds>44,zds,NA)

plot(zds3,d,col='green',main='Zung Depression Scale (ZDS) vs Trauma Symptom Inventory (TSI)',pch=15,xlab='ZDS Level of Depression',ylab='TSI Depression T-score')

y6=lm(d~zds3)

abline(y6,col='black',lwd=3)

summary(y6)

par(mfrow=c(2,2))

plot(y6);ad.test(y6$residuals)

dev.off()

plot(zds2,dm,col='green',main='Zung Depression Scale (ZDS) vs Minnesota Multiphasic Personality Inventory (MMPI)',pch=15,xlab='ZDS Level of Depression',ylab='MMPI Depression T-score')

y7=lm(dm~zds2)

abline(y7,col='black',lwd=3)

summary(y7)

par(mfrow=c(2,2))

plot(y7);ad.test(y7$residuals)

dev.off()

plot(aa,pt,col='blue',main='Trauma Symptom Inventory (TSI) vs Minnesota Multiphasic Personality Inventory (MMPI)',pch=15,xlab='TSI Anxious Arousal T-score',ylab='MMPI Psychathenia T-score')

y8=lm(pt~aa)

abline(y8,col='black',lwd=3)

summary(y8)

par(mfrow=c(2,2))

plot(y8);ad.test(y8$residuals)

dev.off()

plot(d,dm,col='blue',main='Trauma Symptom Inventory (TSI) vs Minnesota Multiphasic Personality Inventory (MMPI)',pch=15,xlab='TSI Depression T-score',ylab='MMPI Depression T-score')

y9=lm(dm~d)

abline(y9,col='black',lwd=3)

summary(y9)

par(mfrow=c(2,2))

plot(y9);ad.test(y9$residuals)

dev.off()

anxiety=data.frame(bai2,aa,pt)

cor(anxiety,use='pairwise.complete.obs')

cor(anxiety,use='pairwise.complete.obs')^2

depression=data.frame(zds2,d,dm)

cor(depression,use='pairwise.complete.obs')

cor(depression,use='pairwise.complete.obs')^2

#Dave Pairwise Comparison (Validity b/t Tests)

plot(inc,vrin,col="blue",main="Inconsistent Response(TSI) vs. Variable Response Inconsistency(MMPI)",pch=1,xlab="Inconsistent Response(inc)",ylab="Variable Response Inconsistency(vrin)")

par(mfrow=c(2,2))

plot(atr,k,col="blue",main="TSI Atypical Response vs MMPI Defensiveness (K)",pch=2,xlab="ATR T-Score",ylab="K T-score")

plot(atr,l,col="red",main="TSI Atypical Response vs. MMPI Lie",pch=3,xlab="ATR T-score",ylab="L T-score")

plot(atr,s,col="green",main="TSI Atypical Response vs MMPI Superlative Self-Presentation",pch=4,xlab="ATR T-Score",ylab="S T-score")

dev.off()

cor(atr,fakegood,use='pairwise.complete.obs')

par(mfrow=c(1,2))

plot(rl,f,col="darkorange",main="TSI Response Level vs MMPI Infrequency (F)",pch=15,xlab="RL T-Score",xlim=c(30,100),ylab="F T-score")

plot(rl,fb,col="green",main="TSI Response Level vs MMPI F Back",pch=16,xlab="RL T-score",xlim=c(30,100),ylab="FB T-score")

dev.off()

fakebad=data.frame(f,fb)

cor(rl,fakebad,use='pairwise.complete.obs')

retention=ifelse(retention<51,retention,NA)

par(mfrow=c(2,2))

plot(aa,retention,col="rosybrown",main="TSI Anxious Arousal vs TOMM Retention",pch=7,xlab="AA T-score",ylab="Retention")

plot(d,retention,col="seagreen",main="TSI Depression vs TOMM Retention",pch=8,xlab="D T-score",ylab="Retention")

plot(ai,retention,col="thistle",main="TSI Anger/Irritability vs TOMM Retention",pch=9,xlab="AI T-score",ylab="Retention")

plot(ie,retention,col="turquoise",main="TSI Intrusive Experiences vs TOMM Retention",pch=10,xlab="IE T-score",ylab="Retention")

dev.off()

par(mfrow=c(2,2))

plot(da,retention,col="navy",main="TSI Defensive Avoidance vs TOMM Retention",pch=11,xlab="DA T-score",ylab="Retention")

plot(dis,retention,col="greenyellow",main="TSI Dissociation vs TOMM Retention ",pch=12,xlab="DIS T-score",ylab="Retention")

plot(sd,retention,col="purple",main="TSI Sexual Concerns vs TOMM Retention",pch=13,xlab="SC T-score",ylab="Retention")

plot(dsb,retention,col="gold",main="TSI Dysfunctional Sexual Behavior vs TOMM Retention",pch=14, xlab="DSB T-score", ylab="Retention")

dev.off()

par(mfrow=c(1,2))

plot(isr,retention,col="dodgerblue",main="TSI Impaired Self-Reference vs TOMM Retention",pch=15,xlab="ISR T-score",ylab="Retention")

plot(trb,retention,col="firebrick",main="TSI Tension Reduction Behavior vs TOMM Retention",pch=16,xlab="TRB T-score",ylab="Retention")

dev.off()

retention2=ifelse(retention<45,retention,NA)

(length(retention)-sum(is.na(retention2)))/sum(!is.na(retention))

tsiclin=data.frame(aa,d,ai,ie,da,dis,sd,dsb,isr,trb)

cor(retention,tsiclin,use='pairwise.complete.obs')

mmpiclin=data.frame(hs,dm,hy,pd,mf,pa,pt,sc,ma,si)

cor(retention,mmpiclin,use='pairwise.complete.obs')

#Allen Multiple/Simple Regression (Validity w/i MMPI)

names(dat)

fakegood=data.frame(k,l,s)

pairs(fakegood)

y10=lm(k~l);y11=lm(k~s);y12=lm(l~s)

summary(y10);summary(y11);summary(y12)

y13=lm(k~l\*s);y14=lm(l~k\*s);y15=lm(s~k\*l)

summary(y13);summary(y14);summary(y15)

y16=lm(s~k+l)

summary(y16)

par(mfrow=c(2,2))

plot(y16);ad.test(y16$residuals)

dev.off()

resid=ifelse(abs(y16$residuals)<15,y16$residuals,NA)

sum(is.na(resid))

par(mfrow=c(1,2))

plot(y16$fitted,resid);abline(h=0);

qqnorm(resid);qqline(resid)

ad.test(resid)

dev.off()

plot(f,fb,col='purple',main='MMPI "Faking Bad"',pch=19,xlab='F T-score (in first half of test)',ylab='FB T-score (in second half of test)')

y17=lm(f~fb)

abline(y17,col='black',lwd=3)

summary(y17)

par(mfrow=c(2,2))

plot(y17);ad.test(y17$residuals)

dev.off()

resid2=ifelse(abs(y17$residuals)<40,y17$residuals,NA)

sum(is.na(resid2))

par(mfrow=c(1,2))

plot(y17$fitted,resid2);abline(h=0);

qqnorm(resid2);qqline(resid2)

ad.test(resid2)

dev.off()

plot(fitted(y17),resid(y17))

y18=lm(f~fb\*I(fb^2))

summary(y18)

par(mfrow=c(2,2))

plot(y18);ad.test(y18$residuals)

dev.off()

resid3=ifelse(abs(y18$residuals)<40,y18$residuals,NA)

sum(is.na(resid3))

par(mfrow=c(1,2))

plot(y18$fitted,resid3);abline(h=0);

qqnorm(resid3);qqline(resid3)

ad.test(resid3)

dev.off()

boxcox(y17)

sf=sqrt(f)

y19=lm(sf~fb)

summary(y19)

par(mfrow=c(2,2))

plot(y19);ad.test(y19$residuals)

dev.off()

dat=read.csv("stt592dat.csv", header=T)

attach(dat)

names(dat)

dat=data.frame(chartn,ptsd,dat[,19:31])

dat=dat[complete.cases(dat),]

dim(dat)

attach(dat)

names(dat)

install.packages('leaps')

library(leaps)

#Logistic Regression

ynull=glm(ptsd~1)

yfull=glm(ptsd~(inc+atr+rl+aa+d+ai+ie+da+dis+sd+dsb+isr+trb)^2,data=dat,family=binomial)

step(ynull,scope=list(lower=ynull,upper=yfull),direction='forward')

#Backward Elimination commented out for run time

#step(yfull,direction='backward')

step(ynull,scope=list(lower=ynull,upper=yfull),direction='both')

y=glm(ptsd~aa\*da,family=binomial)

summary(y)

prob=predict(y,type="response")

pred=rep(F,length(ptsd))

pred[prob>.5]=T

acctab=table(pred,ptsd);acctab

accuracy=sum(diag(acctab))/sum(acctab); accuracy

specificity=acctab[1]/sum(acctab[1:2]);specificity

sensitivity=acctab[4]/sum(acctab[3:4]);sensitivity

y2=glm(ptsd~inc+atr+rl+aa+d+ai+ie+da+dis+sd+dsb+isr+trb+inc\*da+atr\*rl+atr\*aa+atr\*sd+atr\*trb+rl\*ai+rl\*dis+rl\*sd+aa\*ie+aa\*da+d\*ai+ai\*isr+ai\*trb+ie\*dsb+ie\*trb+da\*dis+da\*sd+da\*dsb+da\*trb+dis\*sd+dis\*isr,data=dat,family=binomial)

summary(y2)

prob=predict(y2,type="response")

pred=rep(F,length(ptsd))

pred[prob>.5]=T

acctab=table(pred,ptsd);acctab

accuracy=sum(diag(acctab))/sum(acctab); accuracy

specificity=acctab[1]/sum(acctab[1:2]);specificity

sensitivity=acctab[4]/sum(acctab[3:4]);sensitivity

length(ptsd)\*0.8

dat[493,1]

train=(chartn<31464)

test=dat[!train,]

ptest=ptsd[!train]

ytrain=glm(ptsd~aa\*da,data=dat,family=binomial,subset=train)

summary(ytrain)

ptrain=predict(ytrain,test,type="response")

pred=rep(F,length(ptest))

pred[ptrain>.5]=T

acctab=table(pred,ptest);acctab

accuracy=sum(diag(acctab))/sum(acctab); accuracy

specificity=acctab[1]/sum(acctab[1:2]);specificity

sensitivity=acctab[4]/sum(acctab[3:4]);sensitivity