Investigating a new stability score

Matt Tyers

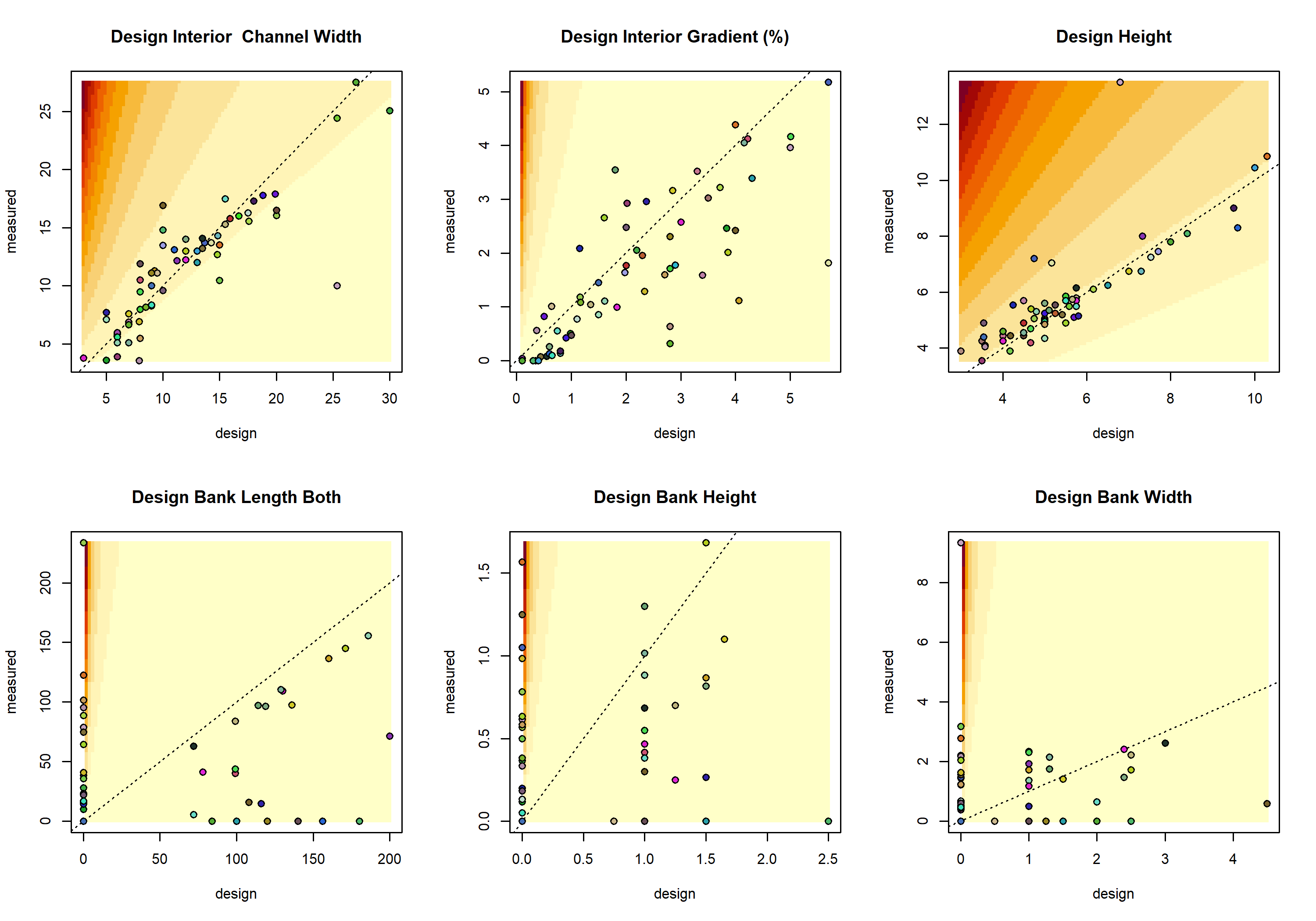
2024-06-13

## The current VTable score

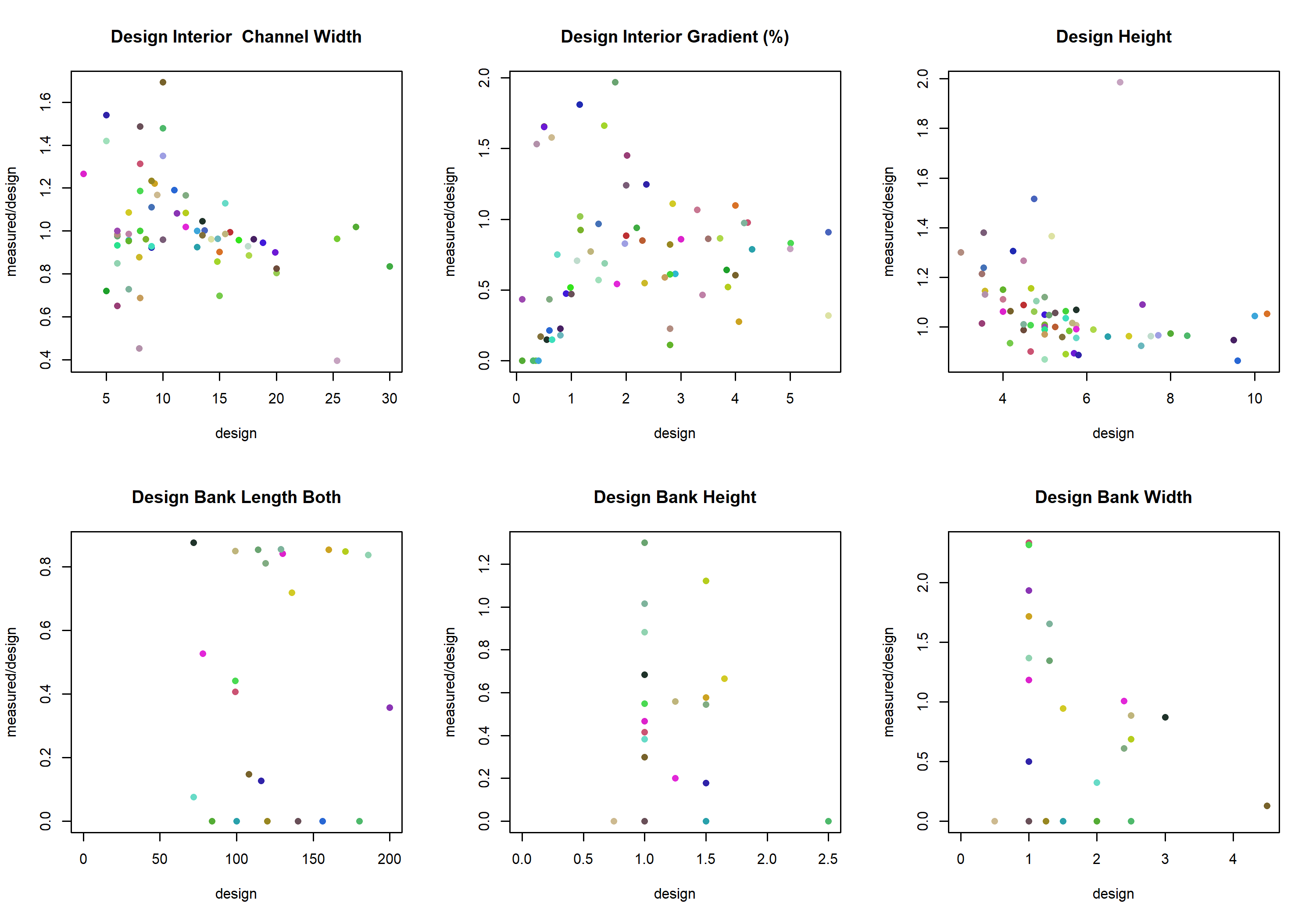
Currently, stability scores are calculated according to . Many design values pertaining to banks were missing; these were imputed to have a value of zero.

It was very useful to me to see the design vs measurement values plotted for each possible variable. Channel width, interior gradient, and height were relatively consistent with one obvious outlier. Channel width and height seemed to vary with additive error (near-constant noise about the y=x line), but gradient appeared to vary with multiplicative error (increasing with x), suggesting that a ratiometric stability score is appropriate. Bank length seemed to be either very consistent or else NOT: the measured value was either very close to the design value or else collapsed to zero (or jumped to a positive value following a zero design). Bank height and bank width were difficult to qualitatively assess!

Scatterplots are shown below for all variables, with an overlayed color ramp depicting a surface generated from the stability score rule (light yellow indicating small values, dark red indicating large values.) The y=x line is overlayed as a dotted line.



The next sequence of scatterplots show the relationships between the design value and calculated V score. I mostly wanted to make sure that there wasn’t an induced effect in which the magnitude of the design value influenced the values the V score could take on (like there would have been if we had subtracted intead of divided.)

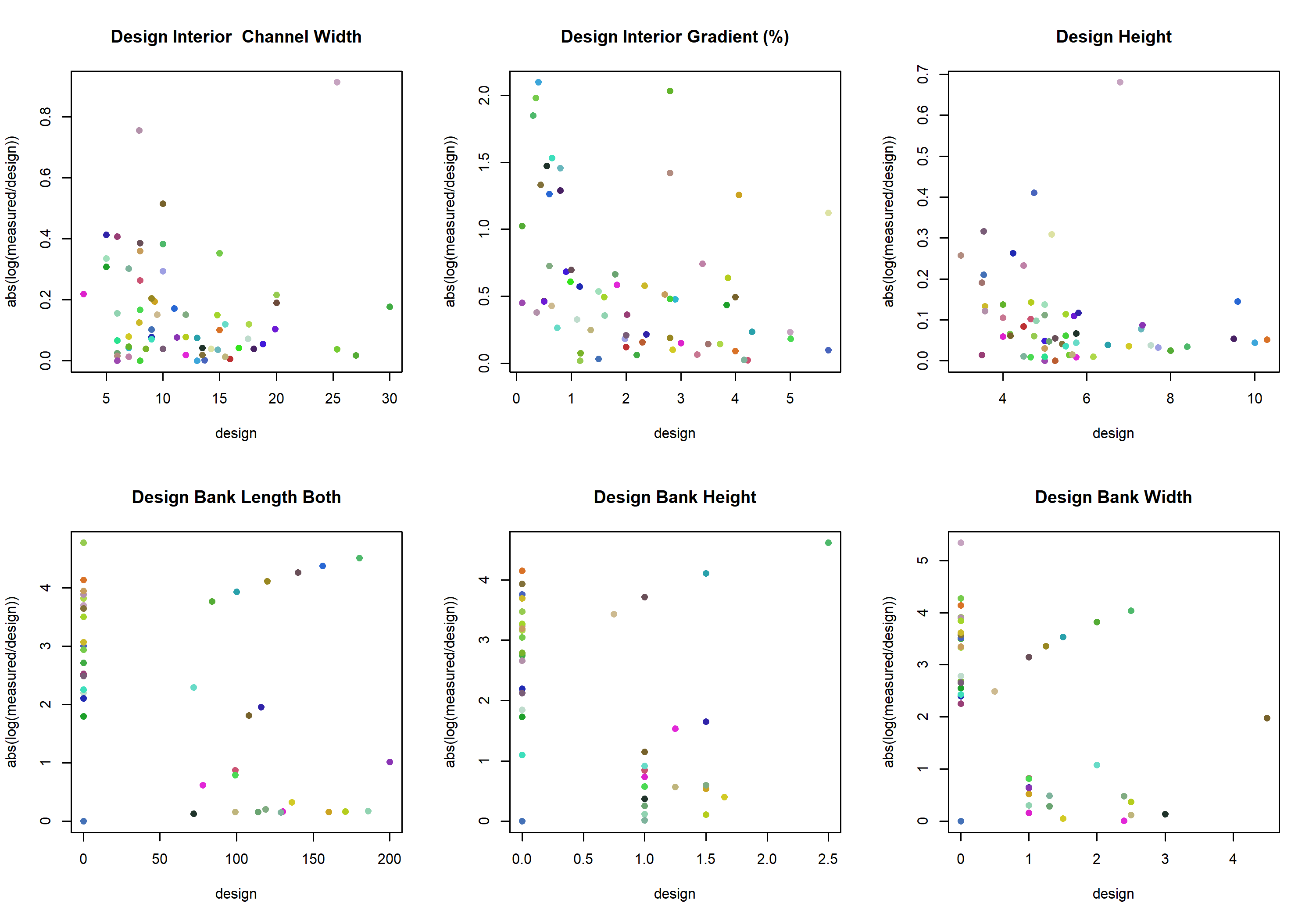
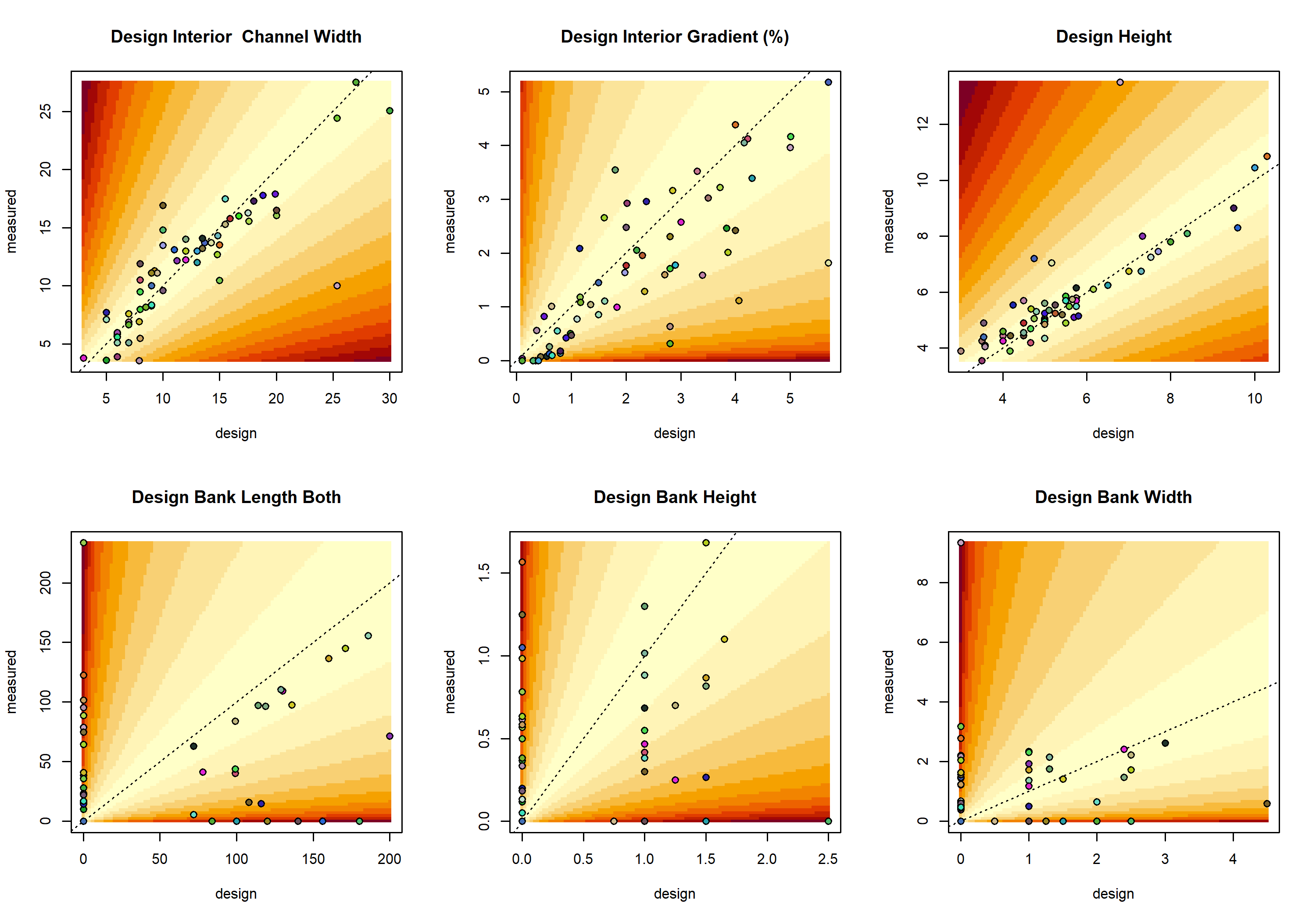


## Proposing a new V score

I would like to propose a new V score to consider: . It looks a little like algebraic soup, but there’s some theoretical underpinning to the ugliness. First, the log transformation has the effect of recasting variability on the multiplicative scale as happening on the additive scale. Functionally, this means that a measured value that is 2x the design value is the same distance from equal as a measured value that is 0.5x the design value. Second, the absolute value forces stable values to have low scores and unstable values to have high scores.

Big note: Since there were values of zero for design or measurement, a small constant was added to both, equal to 1% of the range of the associated design value. This made the log term not become undefined, since it’s no fun when that happens.

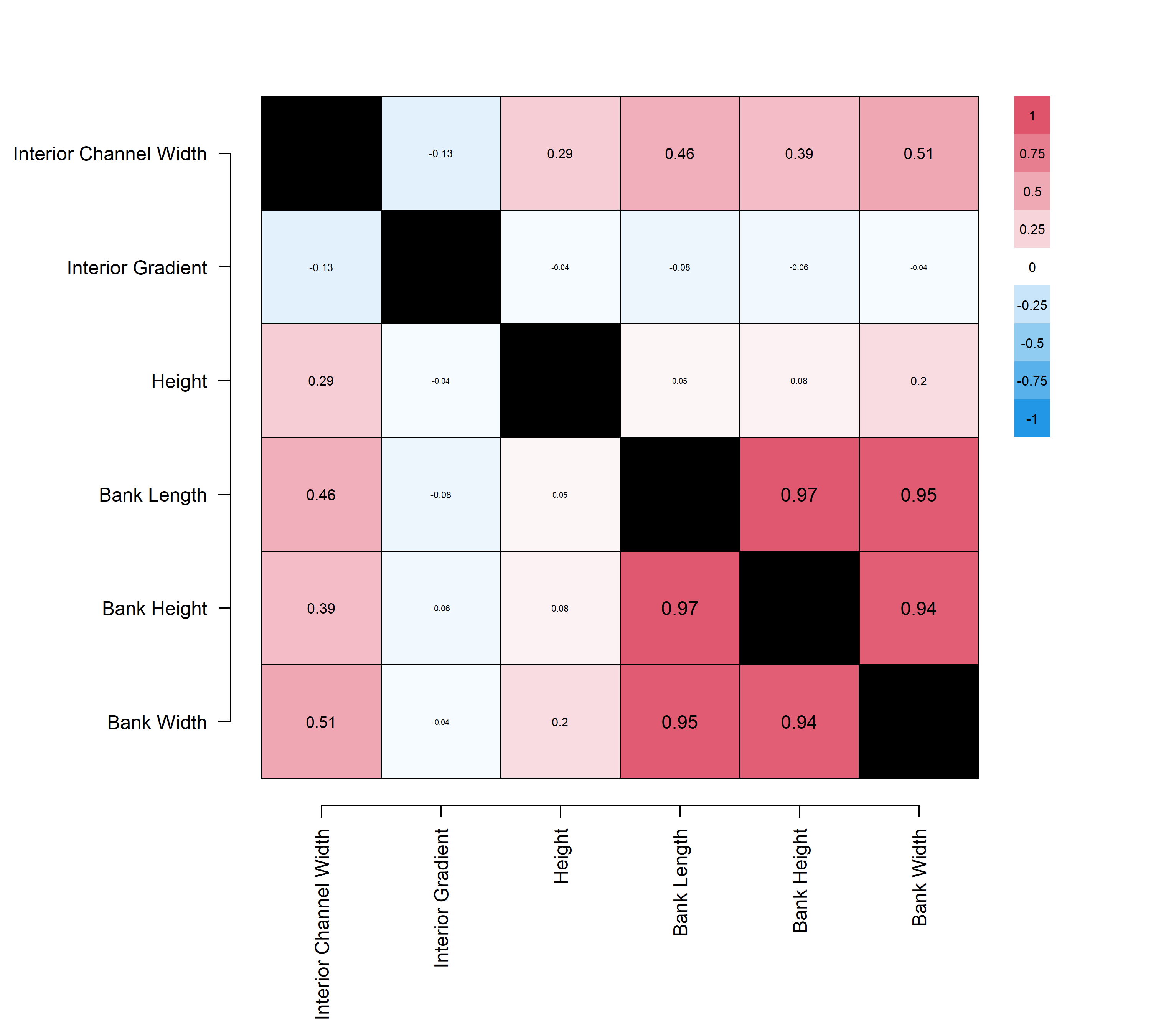
Here are scatterplots for all variables, with an overlayed color ramp associated with the proposed V scoring. Again, points falling in a red zone have a large score; points falling in a yellow zone have a small score.

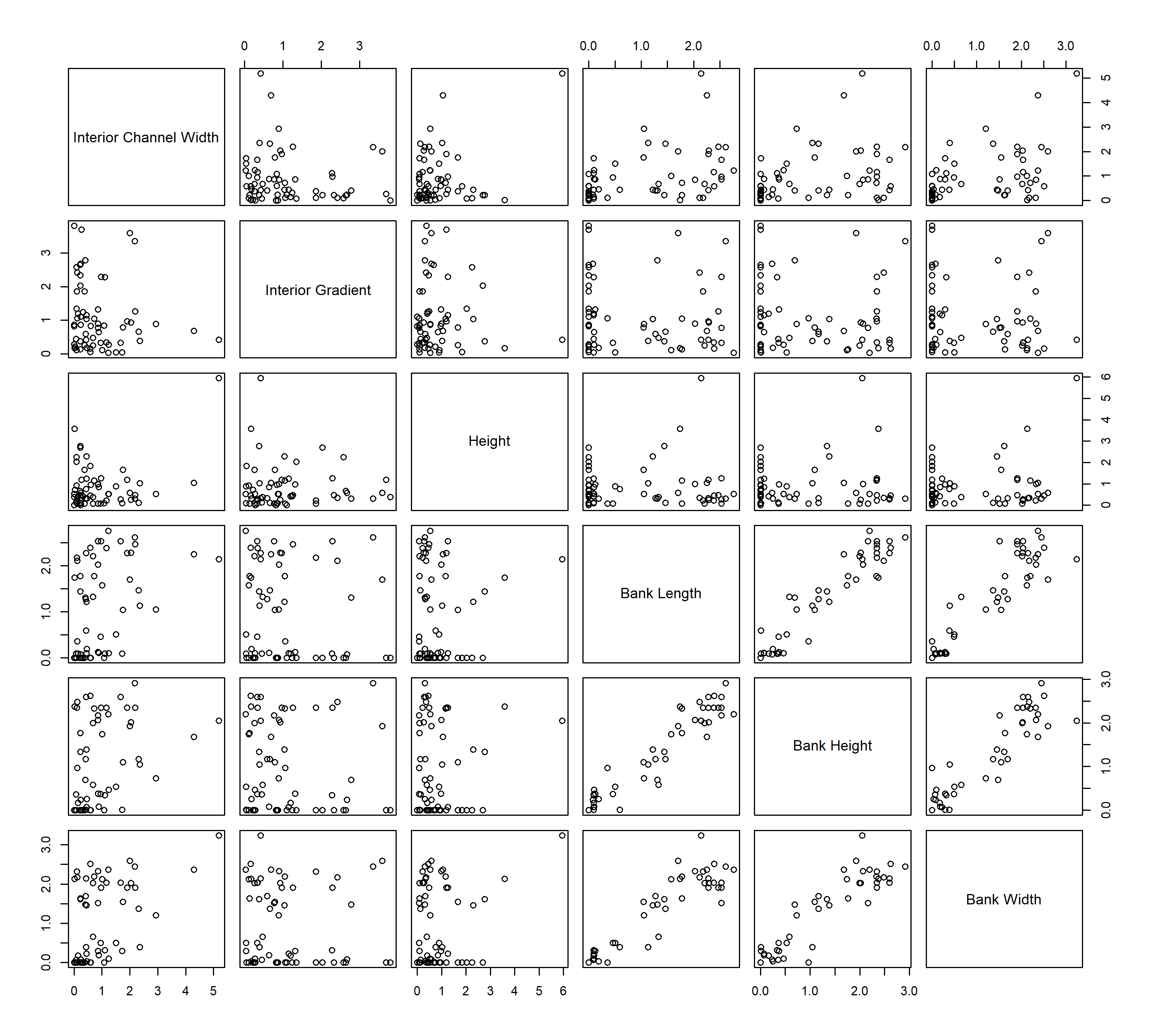


## Correlation between the new V scores

The correlation matrix between V scores is printed below, as are scatterplots between all pairs of scores. Note the high correlation between all the bank values!

|  | Interior Channel Width | Interior Gradient | Height | Bank Length | Bank Height | Bank Width |
| --- | --- | --- | --- | --- | --- | --- |
| Interior Channel Width | 1.000 | -0.132 | 0.291 | 0.464 | 0.388 | 0.515 |
| Interior Gradient | -0.132 | 1.000 | -0.044 | -0.081 | -0.058 | -0.043 |
| Height | 0.291 | -0.044 | 1.000 | 0.051 | 0.079 | 0.203 |
| Bank Length | 0.464 | -0.081 | 0.051 | 1.000 | 0.968 | 0.945 |
| Bank Height | 0.388 | -0.058 | 0.079 | 0.968 | 1.000 | 0.938 |
| Bank Width | 0.515 | -0.043 | 0.203 | 0.945 | 0.938 | 1.000 |





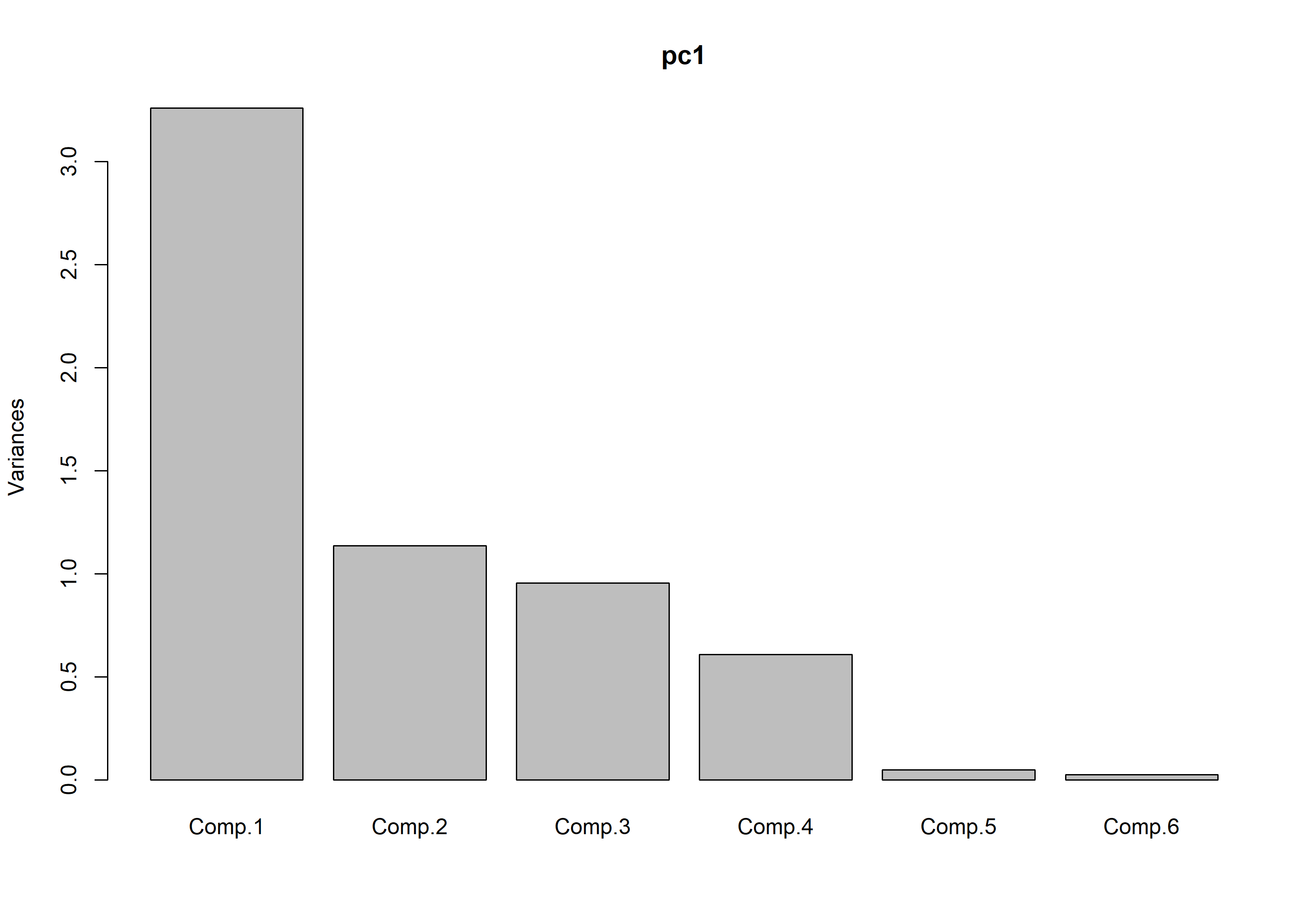
## Aggregating scores?

Since there is a high degree of correlation between the V scores for these six variables, I was curious how much variability could be collapsed to a smaller number of measures.

For this, I tried a Principal Components Analysis, which rotates k-dimensional data such that the maximum amount of variability occurs along a single (aggregate) axis, then such that the maximum amount of remaining variability occurs along an orthogonal axis, and so on. Think of a 2-d scatterplot with a strong correlation, and rotating it such that most of the data can be explained by distance along a single line, and extending this concept to k dimensions (in our case, six).

Note: scores were standardized such that each had unit variance.

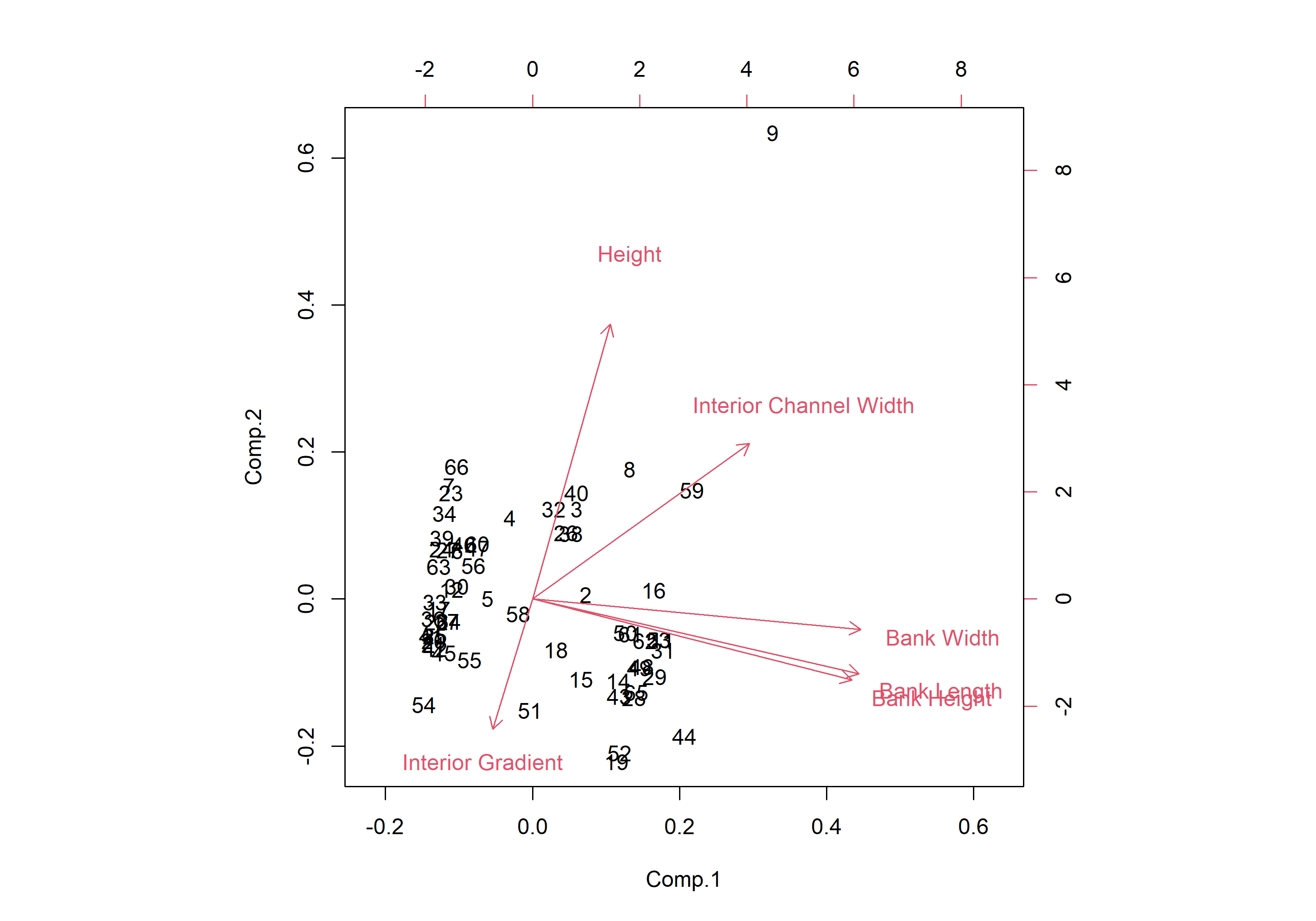
The principal components scoring was as summarized below, suggesting that over half of the variability among the six scores could be explained by a single variable, and that nearly all the variability among the six V scores could be explained by collapsing to four variables (not surprising, given how collinear the three bank scores were.)



## Importance of components:  
## Comp.1 Comp.2 Comp.3 Comp.4 Comp.5  
## Standard deviation 1.8056493 1.0663862 0.9777690 0.7803129 0.221944165  
## Proportion of Variance 0.5399449 0.1883266 0.1583271 0.1008371 0.008157744  
## Cumulative Proportion 0.5399449 0.7282715 0.8865985 0.9874356 0.995593355  
## Comp.6  
## Standard deviation 0.163122076  
## Proportion of Variance 0.004406645  
## Cumulative Proportion 1.000000000

Some additional output is below. PCA is famously difficult to interpret, and I’m now several years out of grad school. I would like to do a bit more research before making an idiot of myself…. however, I *think* that the plot below shows a representation of how correlated the six variables are within the first two principal components. All the Bank variables are highly correlated with one another, and nearly completely independendent of Height and Interior Gradient, which are negatively correlated with one another. Interior Channel width is somewhat correlated with Height and the Bank stuff.

The first PC is mostly made of the three Bank variables, the second is mostly Height, the third is mostly Gradient, and the fourth is mostly Interior Channel width. Since all PC’s are fully orthogonal to one another, it can be suggested that the information in Height, Gradient, and Interior Channel Width is somewhat mutually independent.



##   
## Loadings:  
## Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6  
## Interior Channel Width 0.356 0.432 0.821   
## Interior Gradient -0.362 0.913 0.174   
## Height 0.127 0.763 0.398 -0.480   
## Bank Length 0.536 -0.207 0.298 -0.755  
## Bank Height 0.524 -0.225 -0.211 0.465 0.643  
## Bank Width 0.539 -0.118 -0.823   
##   
## Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6  
## SS loadings 1.000 1.000 1.000 1.000 1.000 1.000  
## Proportion Var 0.167 0.167 0.167 0.167 0.167 0.167  
## Cumulative Var 0.167 0.333 0.500 0.667 0.833 1.000

So what from here? We could possibly use the first PC as our best single metric of instability, since it accounted for over half of the variance among the six scores. Or more conservatively, we use a single PC score to talk about the bank stuff (I just did a PCA with the bank variables by themselves and the first PC explained 97% of the variance!) and interpret the other scores independently of one another. Or, we get fancy and interpret each PC as different *ways* that instability can manifest on the ground. I haven’t had time to think about this yet, but it might be interesting to try.