Evaluating the detectionspace - statistical report

by Milo Marsfeldt Skovfoged, Alexander Schiller Rasmussen

18/5/2020

## Table of Contents

1. [Introduction](#introduction)
2. [Data information](#data-information)
3. [Predicition of course completion time](#predicition-of-course-completion-time)
4. [Collisions and Detections](#collisions-and-detections)

## Introduction

This is the statistical report associated with the paper “Evaluating the detection space” by Milo Marsfeldt Skovfoged & Alexander Schiller rasmussen. The field of study lies in research on behalf of visually impaired/blind navigation through environments, to find an ideal distance and Field of Detection (FOD) for the most progressive travel-route, while avoiding collisions as much as possible. This report gives an overview of what data was gathered and analysed, in regrads to different points of interrest.

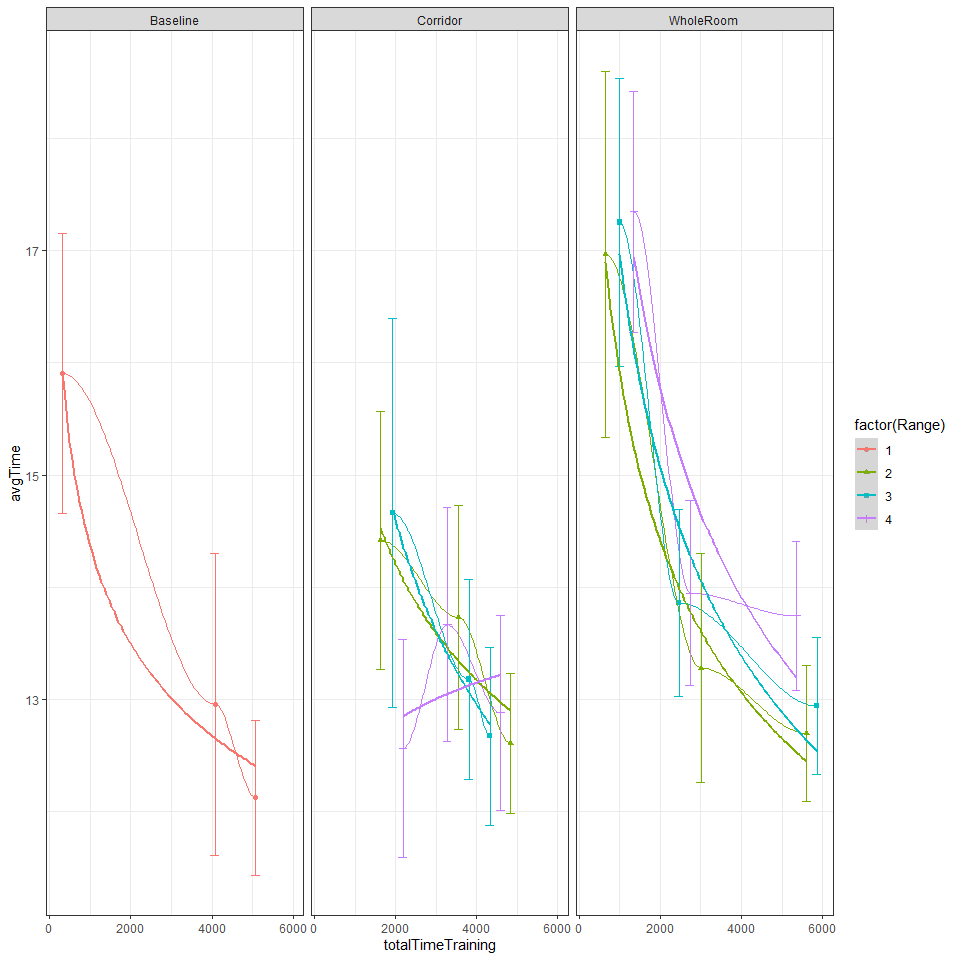
## Data information

Below, a summary of our data is presented. In total, 420 tests were completed over three days (140 per day), using three different Field Of Detections (FOD), with exception of WholeRoom and Corridor differing between three ranges (two, three and four meters). Scenarios describe the amount of different parkours the system were tested on (each condition with different lengths). In addition, each test logged the speed of the participant, the amount of objects detected by the cane, the amount of collisions by the user and the completion time of the individual parkours.

## testID day Scenario FOD Range   
## Min. : 1.0 Min. :1 Min. : 1.00 Baseline : 60 Min. :1.000   
## 1st Qu.:105.8 1st Qu.:1 1st Qu.: 5.75 Corridor :180 1st Qu.:2.000   
## Median :210.5 Median :2 Median :10.50 WholeRoom:180 Median :3.000   
## Mean :210.5 Mean :2 Mean :10.50 Mean :2.714   
## 3rd Qu.:315.2 3rd Qu.:3 3rd Qu.:15.25 3rd Qu.:4.000   
## Max. :420.0 Max. :3 Max. :20.00 Max. :4.000   
## avgSpeed medianSpeed maxSpeed minSpeed objectDetected   
## Min. :0.3595 Min. :0.2511 Min. :1.243 Min. :0 Min. : 0.000   
## 1st Qu.:0.5723 1st Qu.:0.5986 1st Qu.:1.835 1st Qu.:0 1st Qu.: 6.000   
## Median :0.6434 Median :0.6898 Median :2.069 Median :0 Median : 8.000   
## Mean :0.6337 Mean :0.6842 Mean :2.078 Mean :0 Mean : 9.152   
## 3rd Qu.:0.7012 3rd Qu.:0.7587 3rd Qu.:2.318 3rd Qu.:0 3rd Qu.:12.000   
## Max. :0.9803 Max. :1.2506 Max. :2.988 Max. :0 Max. :28.000   
## objectCollisions Time totalTimeTraining timeFDRtrain   
## Min. :0.000 Min. : 8.016 Min. : 18 Min. : 10.00   
## 1st Qu.:0.000 1st Qu.:12.136 1st Qu.:1737 1st Qu.: 77.75   
## Median :1.000 Median :13.625 Median :3154 Median :145.50   
## Mean :1.124 Mean :13.971 Mean :3090 Mean :147.41   
## 3rd Qu.:2.000 3rd Qu.:15.160 3rd Qu.:4522 3rd Qu.:213.25   
## Max. :6.000 Max. :25.143 Max. :5868 Max. :347.00   
## timeFDtrain timeDtrain   
## Min. : 11.0 Min. : 11.0   
## 1st Qu.: 171.5 1st Qu.: 507.5   
## Median : 360.0 Median : 997.5   
## Mean : 391.1 Mean :1004.4   
## 3rd Qu.: 604.0 3rd Qu.:1492.2   
## Max. :1031.0 Max. :2182.0

## Predicition of course completion time

The plot below shows the average time of completion of scenarios, over the course of the total time the user was training, using the system. Inspecting the x-axis, one can tel that the totalTimeTraining seems to differ between the FOD’s. The cause is the order in which the different FOD’s were tested over the three days, which ended up placing the Corridor in the middel of the dataset, leaving the Baseline and WholeRoom with more training time.



To smoothen it out, coefficients for predicting the avg time over training time were calculated, by sorting the aggregated data by FOD and Range - hence the following paramters for estimations were made: **Note: FOD Corridor with a range of 4 meters gives the impression of getting worse over time, which is due to some immediate well achived run on the first day. In addition, the data contains some outliers which will need further investigation.**

###### Baseline data

##   
## Formula: Time ~ b \* totalTimeTraining^z  
##   
## Parameters:  
## Estimate Std. Error t value Pr(>|t|)   
## b 21.45979 1.86920 11.481 < 2e-16 \*\*\*  
## z -0.06329 0.01232 -5.139 3.39e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.437 on 58 degrees of freedom  
##   
## Number of iterations to convergence: 11   
## Achieved convergence tolerance: 8.356e-07

###### Corridor 2 meters data

##   
## Formula: Time ~ b \* totalTimeTraining^z  
##   
## Parameters:  
## Estimate Std. Error t value Pr(>|t|)   
## b 31.32990 9.69222 3.232 0.00203 \*\*  
## z -0.10500 0.03899 -2.693 0.00923 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.026 on 58 degrees of freedom  
##   
## Number of iterations to convergence: 31   
## Achieved convergence tolerance: 4.683e-08

###### Corridor 3 meters data

##   
## Formula: Time ~ b \* totalTimeTraining^z  
##   
## Parameters:  
## Estimate Std. Error t value Pr(>|t|)   
## b 54.72364 27.45127 1.993 0.05092 .   
## z -0.17482 0.06289 -2.780 0.00733 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.557 on 58 degrees of freedom  
##   
## Number of iterations to convergence: 41   
## Achieved convergence tolerance: 1.213e-07

###### Corridor 4 meters data

##   
## Formula: Time ~ b \* totalTimeTraining^z  
##   
## Parameters:  
## Estimate Std. Error t value Pr(>|t|)   
## b 9.81972 5.16585 1.901 0.0623 .  
## z 0.03528 0.06541 0.539 0.5917   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.089 on 58 degrees of freedom  
##   
## Number of iterations to convergence: 32   
## Achieved convergence tolerance: 5.818e-07

###### WholeRoom 2 meters data

##   
## Formula: Time ~ b \* totalTimeTraining^z  
##   
## Parameters:  
## Estimate Std. Error t value Pr(>|t|)   
## b 37.2835 5.6165 6.638 1.18e-08 \*\*\*  
## z -0.1273 0.0204 -6.240 5.46e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.441 on 58 degrees of freedom  
##   
## Number of iterations to convergence: 19   
## Achieved convergence tolerance: 1.844e-07

###### WholeRoom 3 meters data

##   
## Formula: Time ~ b \* totalTimeTraining^z  
##   
## Parameters:  
## Estimate Std. Error t value Pr(>|t|)   
## b 47.50384 8.32368 5.707 4.12e-07 \*\*\*  
## z -0.15325 0.02318 -6.611 1.32e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.087 on 58 degrees of freedom  
##   
## Number of iterations to convergence: 16   
## Achieved convergence tolerance: 2.502e-06

###### WholeRoom 4 meters data

##   
## Formula: Time ~ b \* totalTimeTraining^z  
##   
## Parameters:  
## Estimate Std. Error t value Pr(>|t|)   
## b 54.86090 11.77431 4.659 1.90e-05 \*\*\*  
## z -0.16613 0.02775 -5.986 1.44e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.968 on 58 degrees of freedom  
##   
## Number of iterations to convergence: 27   
## Achieved convergence tolerance: 7.364e-06

Plotting these new predictions onto the data, we see the development in average time of completion for the scenarios over totalTimeTraining. Noteworthy here is that FOD WholeRoom “recieves” a penalty as the Range increases, indicating looking further ahead means slower completion time. Another aspect is that FOD Corridor seems to be at best with a Range of three meters.

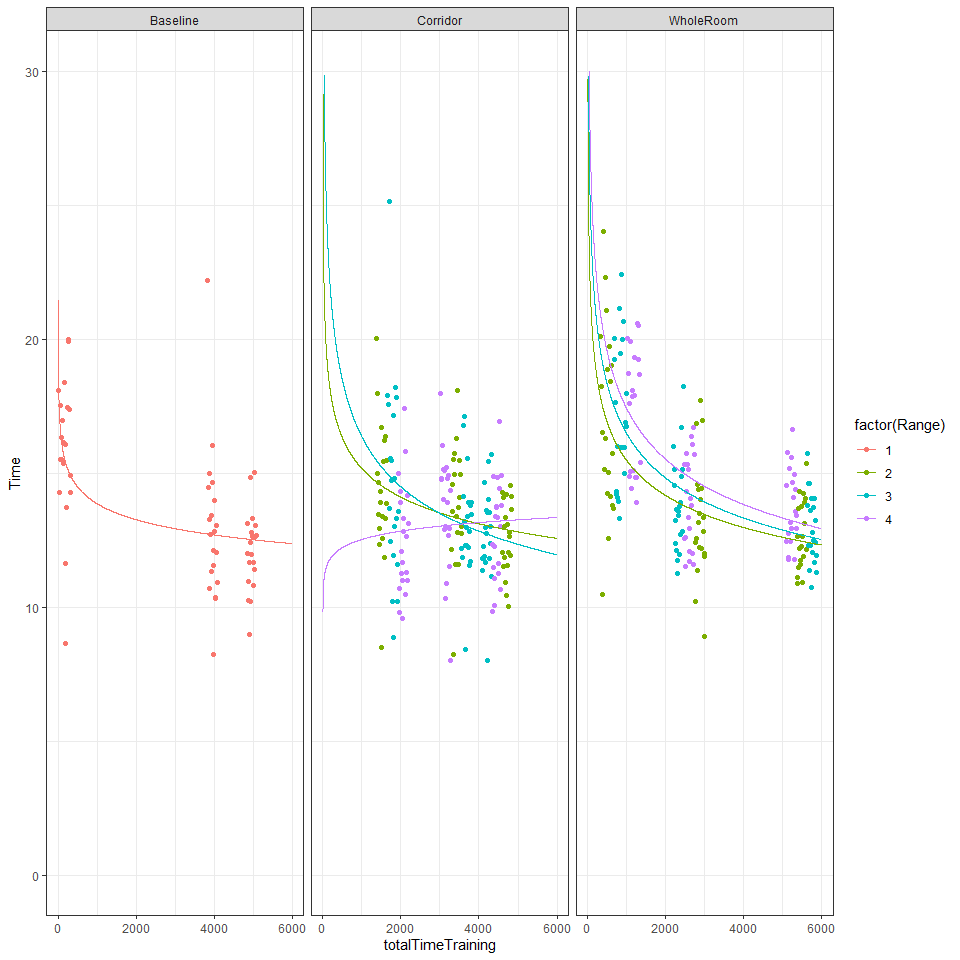
## Warning: Removed 1 row(s) containing missing values (geom\_path).

## Warning: Removed 31 row(s) containing missing values (geom\_path).

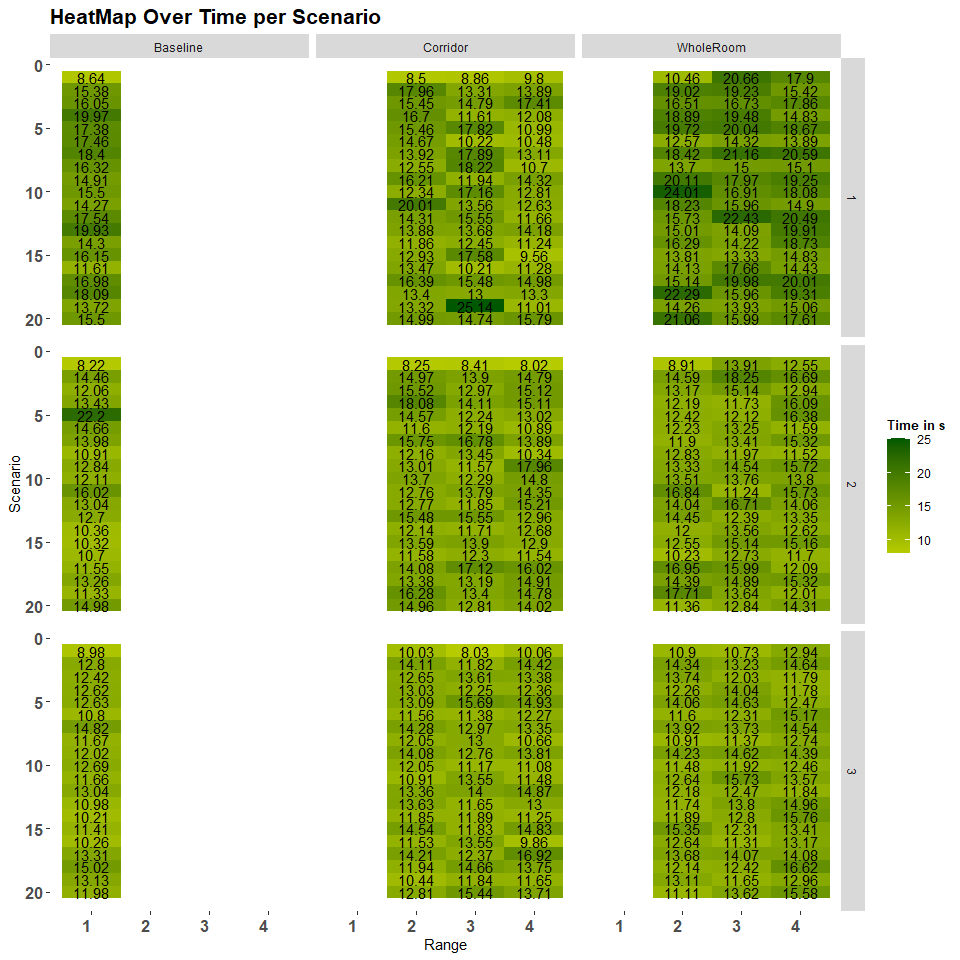
## Warning: Removed 5 row(s) containing missing values (geom\_path).

## Warning: Removed 20 row(s) containing missing values (geom\_path).

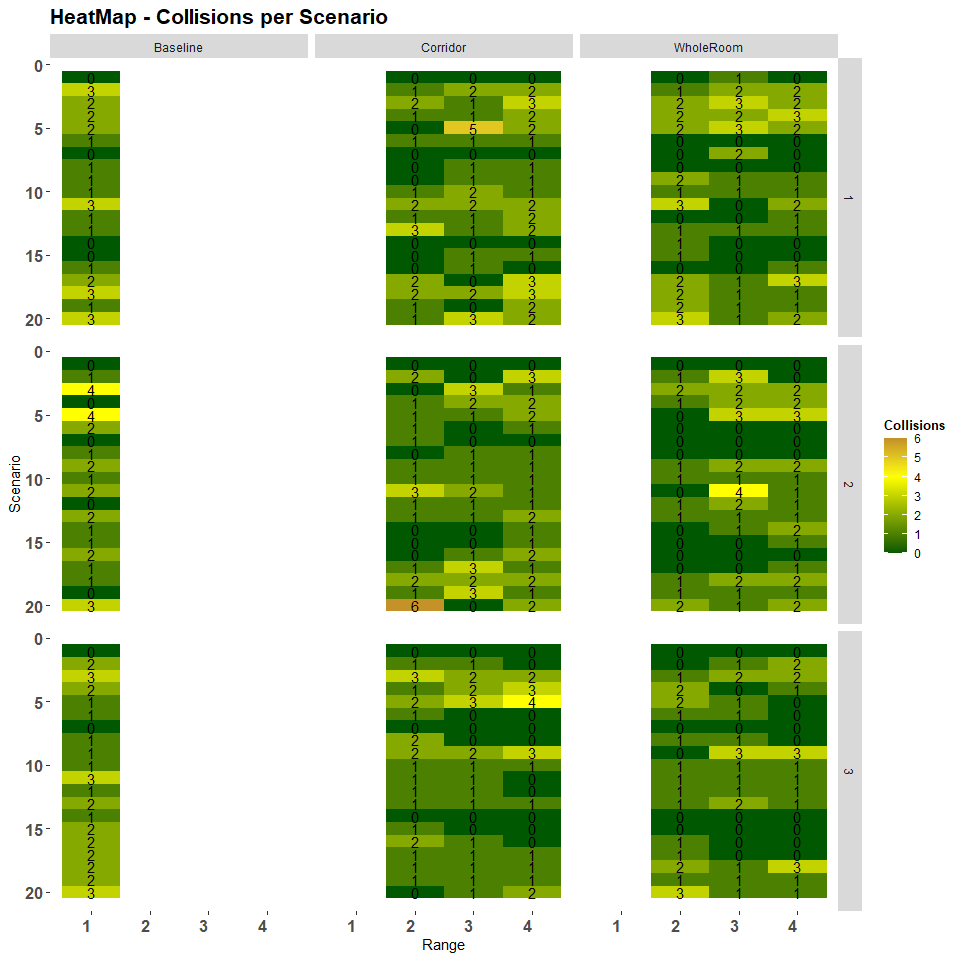
## Warning: Removed 37 row(s) containing missing values (geom\_path).

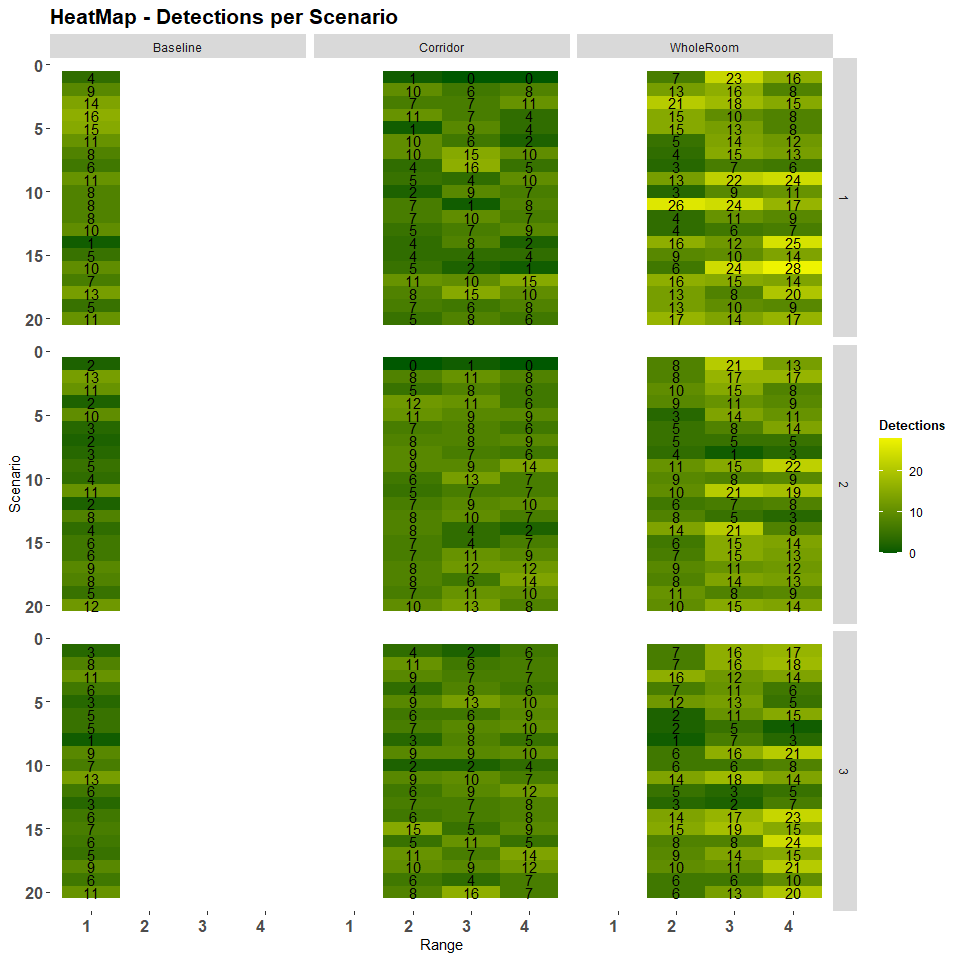


To inspect the time used for completing each scenario, the heatmap shown below plots avg. completion time per scenario for the different FODS with different ranges. Here the outliers stand out, in terms of scenario 19 in DOF Corridor on day one and scenario five in DOF Baseline on day two. Noteworthy is also day one of FOD WholeRoom, which is slower, compared to the other FOD’s.



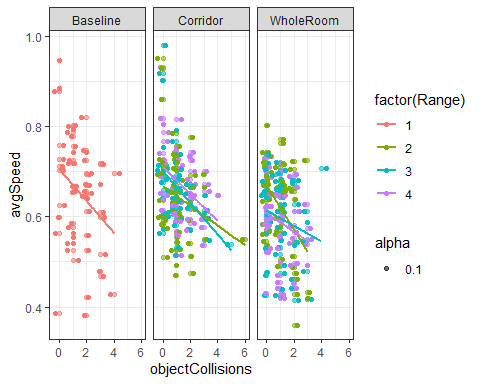
## Collisions and Detections

The heatmap below shows the amount of object collisions for each scenario between, between differnt FOD’s and Ranges. Noteworthy that using WholeRoom DOF, the participant managed to achieve collisions on scenario one, day one, which is supposed to be a clear walkingpath with objects spread out to the sides, making it difficult to collide with an object. 

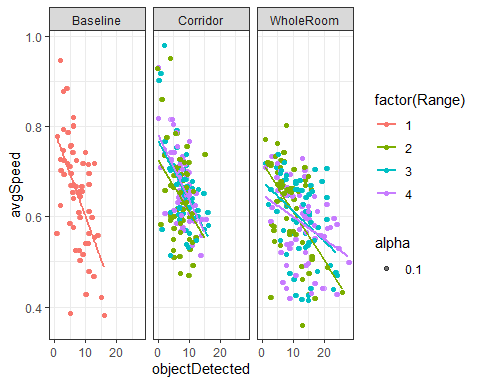
The next heatmap shows the amount of object detections for each scenario, between differnt FOD’s and Ranges. Again, scenario one is very noteworthy for the few detections using FOD Baseline and Corridor, where the walking path is clear, compared to FOD Wholeroom, which conducts alot of information. 

Collisions and detections also have an impact on the speed which the participant moves through scenarios. Below two graphs show that the impact of collisions as well as detections both grow while speed decreases.

ggplot(daggByScen,aes(y = avgSpeed, x = objectCollisions, color = factor(Range)))+  
 geom\_point(aes(alpha=.1))+   
 geom\_jitter()+  
 #geom\_smooth(size=0)+  
 stat\_smooth(method = 'nls', formula = 'y~a+x\*b', method.args = list(start= c(a = 1,b=1)),se=FALSE)+  
 theme\_bw()+  
 facet\_grid(cols=vars(FOD))



ggplot(daggByScen,aes(y = avgSpeed, x = objectDetected, color = factor(Range)))+  
 geom\_point(aes(alpha=.1))+  
 geom\_jitter()+  
 #geom\_smooth(size=0)+  
 stat\_smooth(method = 'nls', formula = 'y~a+x\*b', method.args = list(start= c(a = 1,b=1)),se=FALSE)+  
 theme\_bw()+  
 facet\_grid(cols=vars(FOD))



## Speed

