Believability questionnaire analysis

Zilla Garama

September 2019

Contents

Introduction

This document presents inferential statistical analyses of participants' perception of the believability of virtual characters as reported in the paper:

Creating windows to the soul: How eye gaze behaviour can make virtual characters more believable, more socially present and have mental states attributed to them

The OSF form belonging to this paper can be found here: https://osf.io/8u2kf

Authored by Zilla Garama, Ding Ding, and Willem-Paul Brinkman.

This questionnaire on the perception of the believability of the virtual characters is from P. Gomes, A. Paiva, C. Martinho, and A. Jhala titled "Metrics for character believability in interactive narrative". It was published in International Conference on Interactive Digital Storytelling, pages 223–228, Springer, 2013.
Participants evaluated to what extent each of the sentences matched with the virtual character (VC), using a seven-point Likert scale (1: "Totally disagree", 7: "Totally agree") for each item. The virtual characters have been given a randomised name, as they differ, a placeholder has been used here:

- *Name of the VC* perceives the world around her.
  Totally disagree 1 2 3 4 5 6 7 Totally agree
- It is easy to understand what *Name of the VC* is thinking about.
  Totally disagree 1 2 3 4 5 6 7 Totally agree
- *Name of the VC* has a personality.
  Totally disagree 1 2 3 4 5 6 7 Totally agree
- *Name of the VC*'s behaviour draws my attention.
  Totally disagree 1 2 3 4 5 6 7 Totally agree
- *Name of the VC*'s behaviour is predictable.
  Totally disagree 1 2 3 4 5 6 7 Totally agree
- *Name of the VC*'s behaviour is coherent.
  Totally disagree 1 2 3 4 5 6 7 Totally agree
- *Name of the VC*'s behaviour changes according to experience.
  Totally disagree 1 2 3 4 5 6 7 Totally agree

- *Name of the VC* interacts socially with other characters.
  Totally disagree 1 2 3 4 5 6 7 Totally agree

Abbreviations: * virtual character (VC) * Control virtual character (CVC) * Gaze aware virtual character (GAVC)

Libraries used:

```
library(ggplot2)        # plotting & data
library(psych)          # reliability function
library(pastecs)        # plotting & data
library(lsr)            # effect size
```

Read in the questionnaire data:

```
BQ_data <- read.csv("BQ_data.csv", header = TRUE, stringsAsFactors = FALSE,
                    fileEncoding="UTF-8-BOM")
BQ_data<-BQ_data[-c(25:28),]

#Gaze behaviour data is needed for information about the roles of the virtual characters
GB_data <- read.csv("GB_data.csv", header = TRUE, stringsAsFactors = FALSE,
                    fileEncoding="UTF-8-BOM")
GB_data<-GB_data[-c(25:28),]
```

Data preparation

The participants were asked to fill this questionnaire in for both virtual characters seperately. In the raw data this is represented by A and B for the physical position (left or right) of the virtual characters in the virtual environment during the experiment. However, the roles of the virtual characters (CVC or GAVC) have been randomized so A and B don't directly correspond to the roles of the virtual characters. Before further analysis can be done, information about the roles need to be added to a data table:

```
#calculate the roles per virtual character:
CVC_position = ifelse(GB_data$GB_1_CVC %in% c("UMA_2", "UMA_3S", "UMA_4", "UMA_5S"),
                      "A", "B")
GAVC_position = ifelse(CVC_position %in% c("A"), "B", "A")

#Put the roles in a table to get a clearer overview:
GB_roles <- data.frame(Participant_number = GB_data$Participant_number,
                       CVC = GB_data$GB_1_CVC, GAVC = GB_data$GB_2_GAVC,
                       CVC_role = CVC_position, GAVC_role = GAVC_position)

print(GB_roles)
```

```
##    Participant_number    CVC   GAVC CVC_role GAVC_role
## 1                 180  UMA_4  UMA_5        A         B
## 2                 122 UMA_5S UMA_4S        A         B
## 3                 198  UMA_4  UMA_5        A         B
## 4                 182  UMA_5  UMA_4        B         A
## 5                 132 UMA_4S UMA_5S        B         A
## 6                 107  UMA_2  UMA_3        A         B
## 7                 184  UMA_3  UMA_2        B         A
## 8                 137 UMA_4S UMA_5S        B         A
## 9                 103  UMA_3  UMA_2        B         A
## 10                136  UMA_3  UMA_2        B         A
## 11                128  UMA_4  UMA_5        A         B
```

```
## 12                185 UMA_3S UMA_2S        A         B
## 13                163 UMA_2S UMA_3S        B         A
## 14                112  UMA_3  UMA_2        B         A
## 15                149 UMA_5S UMA_4S        A         B
## 16                173  UMA_3  UMA_2        B         A
## 17                195  UMA_2  UMA_3        A         B
## 18                115  UMA_4  UMA_5        A         B
## 19                169 UMA_4S UMA_5S        B         A
## 20                120 UMA_5S UMA_4S        A         B
## 21                146 UMA_5S UMA_4S        A         B
## 22                178 UMA_2S UMA_3S        B         A
## 23                119 UMA_3S UMA_2S        A         B
## 24                157 UMA_2S UMA_3S        B         A
```

Using the above mapping of roles with A/B we can seperate the questionnaire answers per role. Thus the data table of the believability questionnaire for the CVC is:

```r
BQ_CVC <- data.frame(BQ_1_awareness=rep(NA,24), BQ_2_behaviour_understandability=rep(NA, 24),
                     BQ_3_personality=rep(NA, 24), BQ_4_visual_impact=rep(NA, 24),
                     BQ_5_predictability=rep(NA, 24), BQ_6_behaviour_coherence=rep(NA, 24),
                     BQ_7_change_with_experience=rep(NA, 24), BQ_8_social=rep(NA, 24),
                     stringsAsFactors = FALSE)

for(i in 1:24){
  if(CVC_position[i] == "A"){
    BQ_CVC[i,] <- BQ_data[i, 2:9]
  }
  else{
    BQ_CVC[i,] <- BQ_data[i, 10:17]
  }
}
BQ_CVC <- cbind(Participant_number = BQ_data$Participant_number, BQ_CVC)

print(BQ_CVC)
```

```
##    Participant_number BQ_1_awareness BQ_2_behaviour_understandability
## 1                 180              3                                1
## 2                 122              1                                1
## 3                 198              6                                4
## 4                 182              6                                3
## 5                 132              6                                2
## 6                 107              7                                4
## 7                 184              2                                2
## 8                 137              1                                2
## 9                 103              6                                2
## 10                136              4                                2
## 11                128              3                                5
## 12                185              5                                5
## 13                163              4                                1
## 14                112              5                                3
## 15                149              5                                4
## 16                173              6                                2
## 17                195              6                                2
## 18                115              6                                2
## 19                169              2                                1
```

```
## 20               120              2                                  2
## 21               146              5                                  5
## 22               178              6                                  3
## 23               119              1                                  5
## 24               157              3                                  2
##    BQ_3_personality BQ_4_visual_impact BQ_5_predictability
## 1                 5                  3                   7
## 2                 1                  1                   1
## 3                 5                  7                   7
## 4                 5                  7                   2
## 5                 3                  5                   7
## 6                 3                  5                   4
## 7                 1                  7                   4
## 8                 3                  2                   6
## 9                 4                  1                   3
## 10                5                  4                   5
## 11                3                  4                   7
## 12                6                  5                   4
## 13                1                  3                   5
## 14                5                  4                   3
## 15                6                  7                   6
## 16                4                  3                   4
## 17                6                  4                   6
## 18                2                  2                   4
## 19                2                  6                   5
## 20                4                  2                   2
## 21                4                  6                   6
## 22                3                  6                   7
## 23                2                  5                   7
## 24                5                  1                   3
##    BQ_6_behaviour_coherence BQ_7_change_with_experience BQ_8_social
## 1                         7                           7           3
## 2                         1                           1           1
## 3                         7                           5           2
## 4                         3                           5           3
## 5                         5                           2           1
## 6                         4                           2           2
## 7                         1                           4           1
## 8                         6                           2           1
## 9                         4                           2           1
## 10                        5                           2           1
## 11                        3                           2           2
## 12                        6                           6           6
## 13                        4                           5           2
## 14                        4                           5           4
## 15                        3                           3           1
## 16                        5                           3           2
## 17                        5                           5           4
## 18                        3                           3           2
## 19                        3                           1           1
## 20                        2                           5           1
## 21                        7                           3           2
## 22                        7                           2           2
## 23                        5                           1           6
```

```
## 24                          5                            5            4
```

And the GAVC:

```
BQ_GAVC <- data.frame(BQ_1_awareness=rep(NA,24), BQ_2_behaviour_understandability=rep(NA, 24),
                      BQ_3_personality=rep(NA, 24), BQ_4_visual_impact=rep(NA, 24),
                      BQ_5_predictability=rep(NA, 24), BQ_6_behaviour_coherence=rep(NA, 24),
                      BQ_7_change_with_experience=rep(NA, 24), BQ_8_social=rep(NA, 24),
                      stringsAsFactors = FALSE)

for(i in 1:24){
  if(GAVC_position[i] == "A"){
    BQ_GAVC[i,] <- BQ_data[i, 2:9]
  }
  else{
    BQ_GAVC[i,] <- BQ_data[i, 10:17]
  }
}
BQ_GAVC <- cbind(Participant_number = BQ_data$Participant_number, BQ_GAVC)

print(BQ_GAVC)
```

```
##    Participant_number BQ_1_awareness BQ_2_behaviour_understandability
## 1                 180              3                                1
## 2                 122              3                                1
## 3                 198              7                                3
## 4                 182              3                                6
## 5                 132              6                                2
## 6                 107              5                                4
## 7                 184              5                                3
## 8                 137              6                                3
## 9                 103              6                                3
## 10                136              5                                2
## 11                128              6                                4
## 12                185              3                                2
## 13                163              5                                1
## 14                112              5                                4
## 15                149              4                                3
## 16                173              5                                2
## 17                195              4                                2
## 18                115              5                                2
## 19                169              2                                6
## 20                120              5                                2
## 21                146              4                                5
## 22                178              6                                3
## 23                119              3                                5
## 24                157              6                                2
##    BQ_3_personality BQ_4_visual_impact BQ_5_predictability
## 1                 5                  3                   7
## 2                 1                  5                   1
## 3                 5                  7                   5
## 4                 3                  5                   6
## 5                 4                  6                   7
## 6                 2                  4                   4
## 7                 5                  1                   7
```

```
## 8                  5                  6                  6
## 9                  5                  1                  5
## 10                 5                  4                  4
## 11                 4                  4                  6
## 12                 4                  4                  6
## 13                 1                  4                  5
## 14                 4                  6                  6
## 15                 3                  7                  6
## 16                 4                  3                  4
## 17                 4                  5                  6
## 18                 5                  5                  3
## 19                 3                  5                  7
## 20                 4                  2                  2
## 21                 4                  6                  6
## 22                 3                  3                  6
## 23                 5                  3                  7
## 24                 5                  7                  5
##    BQ_6_behaviour_coherence BQ_7_change_with_experience BQ_8_social
## 1                         7                           7           3
## 2                         4                           1           6
## 3                         6                           6           6
## 4                         6                           2           6
## 5                         5                           2           1
## 6                         4                           2           2
## 7                         6                           3           5
## 8                         6                           6           4
## 9                         6                           2           1
## 10                        5                           2           5
## 11                        4                           4           5
## 12                        6                           4           3
## 13                        3                           1           4
## 14                        6                           3           2
## 15                        4                           6           2
## 16                        3                           3           2
## 17                        6                           3           3
## 18                        3                           5           5
## 19                        6                           1           5
## 20                        2                           2           1
## 21                        7                           3           5
## 22                        7                           1           2
## 23                        6                           1           6
## 24                        7                           1           7
```

Reliability score

The reliability score for the CVC is:

```
alpha(BQ_CVC)
```

```
##
## Reliability analysis
## Call: alpha(x = BQ_CVC)
##
##   raw_alpha std.alpha G6(smc) average_r S/N  ase mean  sd median_r
```

```
##       0.25      0.75    0.86      0.25   3 0.07   20 3.9     0.27
##
##  lower alpha upper     95% confidence boundaries
## 0.12 0.25 0.39
##
##  Reliability if an item is dropped:
##                                 raw_alpha std.alpha G6(smc) average_r S/N
## Participant_number                   0.72      0.73    0.82      0.25 2.6
## BQ_1_awareness                       0.24      0.74    0.86      0.27 2.9
## BQ_2_behaviour_understandability     0.26      0.73    0.82      0.26 2.8
## BQ_3_personality                     0.23      0.71    0.84      0.23 2.5
## BQ_4_visual_impact                   0.21      0.73    0.81      0.25 2.7
## BQ_5_predictability                  0.23      0.74    0.83      0.26 2.9
## BQ_6_behaviour_coherence             0.22      0.70    0.81      0.23 2.4
## BQ_7_change_with_experience          0.21      0.74    0.84      0.26 2.9
## BQ_8_social                          0.24      0.73    0.84      0.26 2.8
##                                 alpha se var.r med.r
## Participant_number                 0.089 0.035  0.27
## BQ_1_awareness                     0.064 0.039  0.29
## BQ_2_behaviour_understandability    0.062 0.033  0.27
## BQ_3_personality                   0.066 0.037  0.24
## BQ_4_visual_impact                 0.072 0.035  0.28
## BQ_5_predictability                0.068 0.030  0.28
## BQ_6_behaviour_coherence           0.065 0.036  0.24
## BQ_7_change_with_experience        0.074 0.027  0.28
## BQ_8_social                        0.067 0.037  0.27
##
##  Item statistics
##                                   n raw.r std.r r.cor r.drop  mean   sd
## Participant_number               24 0.979  0.60  0.58  0.480 149.6 30.3
## BQ_1_awareness                   24 0.241  0.51  0.41  0.188   4.2  1.9
## BQ_2_behaviour_understandability 24 0.083  0.56  0.53  0.044   2.7  1.4
## BQ_3_personality                 24 0.409  0.68  0.63  0.369   3.7  1.6
## BQ_4_visual_impact               24 0.541  0.57  0.55  0.498   4.2  2.0
## BQ_5_predictability              24 0.358  0.52  0.48  0.310   4.8  1.8
## BQ_6_behaviour_coherence         24 0.455  0.70  0.69  0.413   4.4  1.8
## BQ_7_change_with_experience      24 0.555  0.52  0.47  0.518   3.4  1.7
## BQ_8_social                      24 0.293  0.55  0.49  0.253   2.3  1.5
```

The reliability score for the GAVC is:

```r
alpha(BQ_GAVC)
```

```
## Warning in alpha(BQ_GAVC): Some items were negatively correlated with the total scale and probably
## should be reversed.
## To do this, run the function again with the 'check.keys=TRUE' option
```

```
## Some items ( BQ_1_awareness ) were negatively correlated with the total scale and
## probably should be reversed.
## To do this, run the function again with the 'check.keys=TRUE' option
```

```
##
## Reliability analysis
## Call: alpha(x = BQ_GAVC)
##
##   raw_alpha std.alpha G6(smc) average_r S/N   ase mean  sd median_r
```

```
##       0.14      0.58    0.73      0.13 1.4 0.071   20 3.6     0.11
##
##  lower alpha upper     95% confidence boundaries
## 0 0.14 0.28
##
##  Reliability if an item is dropped:
##                                  raw_alpha std.alpha G6(smc) average_r
## Participant_number                   0.539      0.55    0.69     0.131
## BQ_1_awareness                       0.153      0.64    0.75     0.183
## BQ_2_behaviour_understandability     0.142      0.58    0.71     0.148
## BQ_3_personality                     0.132      0.52    0.64     0.119
## BQ_4_visual_impact                   0.127      0.56    0.69     0.137
## BQ_5_predictability                  0.096      0.48    0.63     0.102
## BQ_6_behaviour_coherence             0.099      0.46    0.63     0.096
## BQ_7_change_with_experience          0.114      0.56    0.68     0.136
## BQ_8_social                          0.114      0.57    0.70     0.141
##                                   S/N alpha se var.r med.r
## Participant_number               1.21    0.142 0.048 0.115
## BQ_1_awareness                   1.79    0.069 0.037 0.194
## BQ_2_behaviour_understandability 1.39    0.066 0.041 0.135
## BQ_3_personality                 1.08    0.065 0.048 0.115
## BQ_4_visual_impact               1.27    0.065 0.054 0.150
## BQ_5_predictability              0.91    0.068 0.037 0.097
## BQ_6_behaviour_coherence         0.85    0.067 0.038 0.084
## BQ_7_change_with_experience      1.26    0.068 0.049 0.097
## BQ_8_social                      1.32    0.069 0.049 0.115
##
##  Item statistics
##                                   n  raw.r std.r  r.cor r.drop  mean   sd
## Participant_number               24  0.984  0.49  0.404  0.303 149.6 30.3
## BQ_1_awareness                   24 -0.119  0.15 -0.019 -0.158   4.7  1.3
## BQ_2_behaviour_understandability 24  0.023  0.38  0.285 -0.022   3.0  1.5
## BQ_3_personality                 24  0.139  0.57  0.537  0.102   3.9  1.2
## BQ_4_visual_impact               24  0.177  0.45  0.347  0.124   4.4  1.7
## BQ_5_predictability              24  0.483  0.68  0.688  0.444   5.3  1.6
## BQ_6_behaviour_coherence         24  0.480  0.72  0.727  0.444   5.2  1.5
## BQ_7_change_with_experience      24  0.282  0.46  0.383  0.229   3.0  1.9
## BQ_8_social                      24  0.283  0.42  0.324  0.229   3.8  1.9
```

The raw alpha reliability score is 0.72 for the CVC and 0.54 for the GAVC. These values are considered acceptable to use the averages of the items in subsequent analyses.

Can individual items be removed?

If we look at the individual item reliability if a single item was removed from the questionnaire. The max score for the CVC would be 0.72 if BQ_7_change with experience was deleted. This is the same reliability score as when no items were deleted. For the GAVC the max reliability score would be 0.59 if BQ_1_awareness was deleted. This is only slightly higher (0.05 higher) then if no items were deleted. Considering that there would be almost no improvement if individual items were removed from the questionnaire and that the item to be removed would differ in both conditions, the questionnaire will be considered in its entirety and no items will be removed.

Combining into a single score

As the reliability score was considered acceptable, a single score for CVC and GAVC is calculated by averaging the individual items of the questionnaire:

```
BQ_CVC_average = ifelse(CVC_position %in% c("A"), rowMeans(BQ_data[2:9]),
                        rowMeans(BQ_data[10:17]))
BQ_GAVC_average = ifelse(GAVC_position %in% c("A"), rowMeans(BQ_data[2:9]),
                         rowMeans(BQ_data[10:17]))

BQ_averages <- data.frame(Participant_number = BQ_data$Participant_number,
                          CVC_average_score = BQ_CVC_average,
                          GAVC_average_score = BQ_GAVC_average)

print(BQ_averages)
```

```
##    Participant_number CVC_average_score GAVC_average_score
## 1                 180             4.500              4.500
## 2                 122             1.000              2.750
## 3                 198             5.375              5.625
## 4                 182             4.250              4.625
## 5                 132             3.875              4.125
## 6                 107             3.875              3.375
## 7                 184             2.750              4.375
## 8                 137             2.875              5.250
## 9                 103             2.875              3.625
## 10                136             3.500              4.000
## 11                128             3.625              4.625
## 12                185             5.375              4.000
## 13                163             3.125              3.000
## 14                112             4.125              4.500
## 15                149             4.375              4.375
## 16                173             3.625              3.250
## 17                195             4.750              4.125
## 18                115             3.000              4.125
## 19                169             2.625              4.375
## 20                120             2.500              2.500
## 21                146             4.750              5.000
## 22                178             4.500              3.875
## 23                119             4.000              4.500
## 24                157             3.500              5.000
```

Assumption checking: normal distribution

The analysis method depends on the normality of the data distribution. This is usually done visually:

```
stem(BQ_averages$CVC_average_score)
```

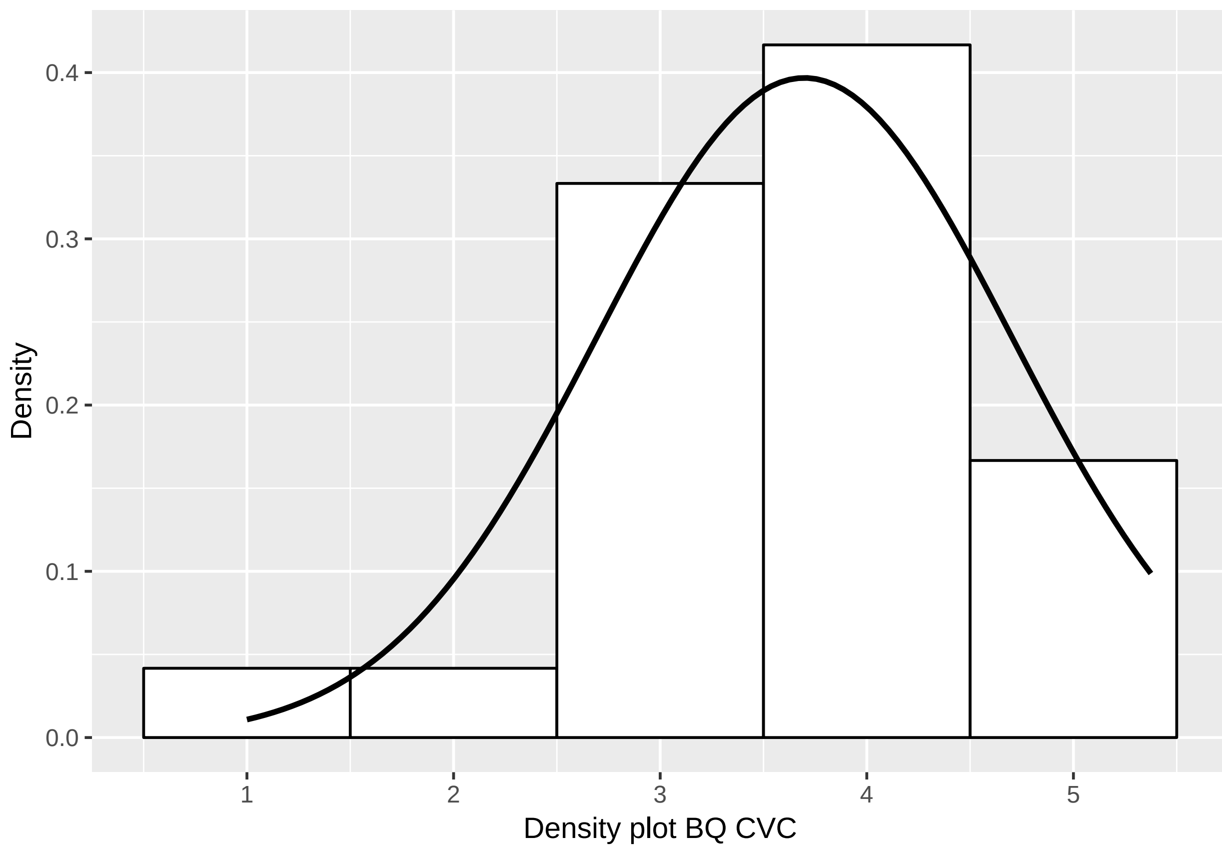
```
##
##   The decimal point is at the |
##
##   1 | 0
##   2 | 56899
##   3 | 01556699
##   4 | 01345588
```

```
##   5 | 44
```

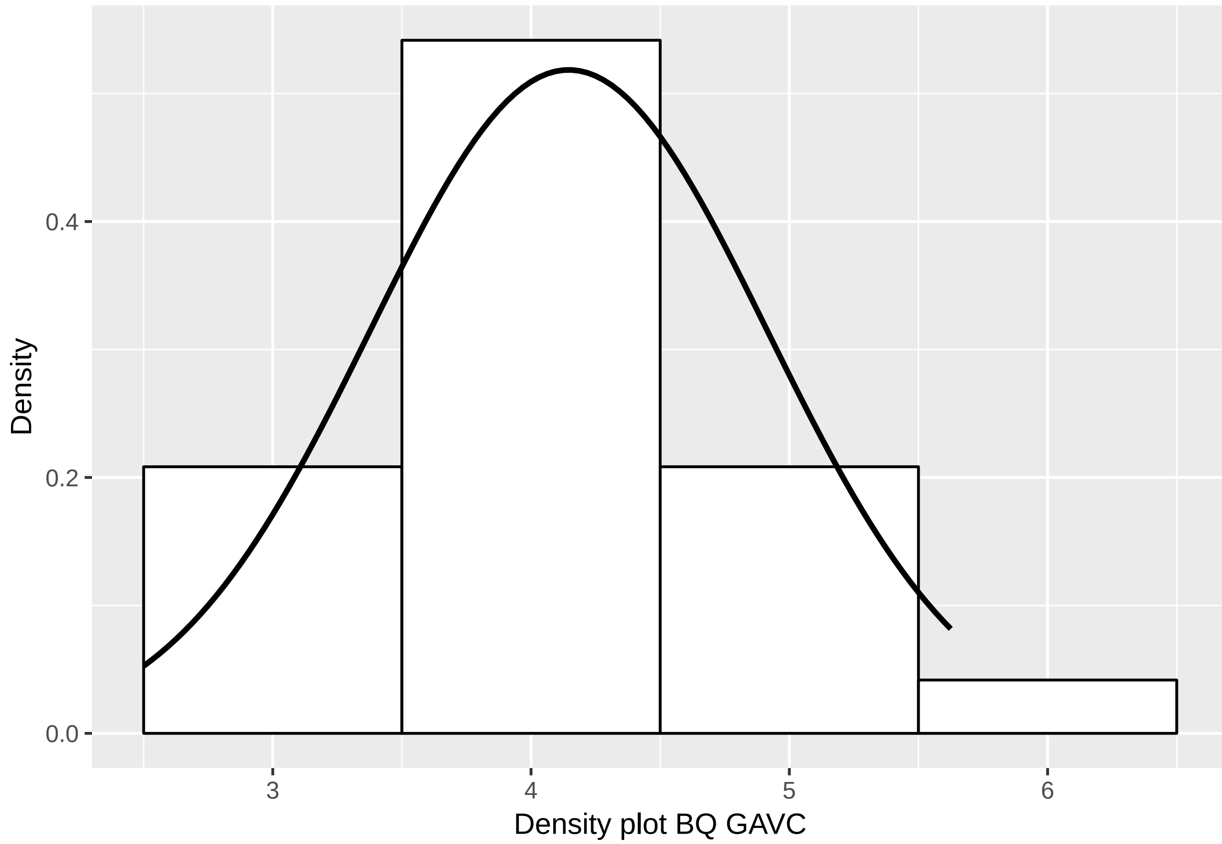
```r
stem(BQ_averages$GAVC_average_score)
```

```
##
##   The decimal point is at the |
##
##   2 | 58
##   3 | 03469
##   4 | 0011144455566
##   5 | 0036
```

```r
ggplot(BQ_averages,
       aes(CVC_average_score)) + geom_histogram(aes(y=..density..),binwidth = 1,
       colour="black", fill="white") + labs(x="Density plot BQ CVC",
       y="Density") + stat_function(fun=dnorm,
       args=list(mean=mean(BQ_averages$CVC_average_score, na.rm=TRUE),
       sd=sd(BQ_averages$CVC_average_score, na.rm=TRUE)), colour="black", size=1)
```

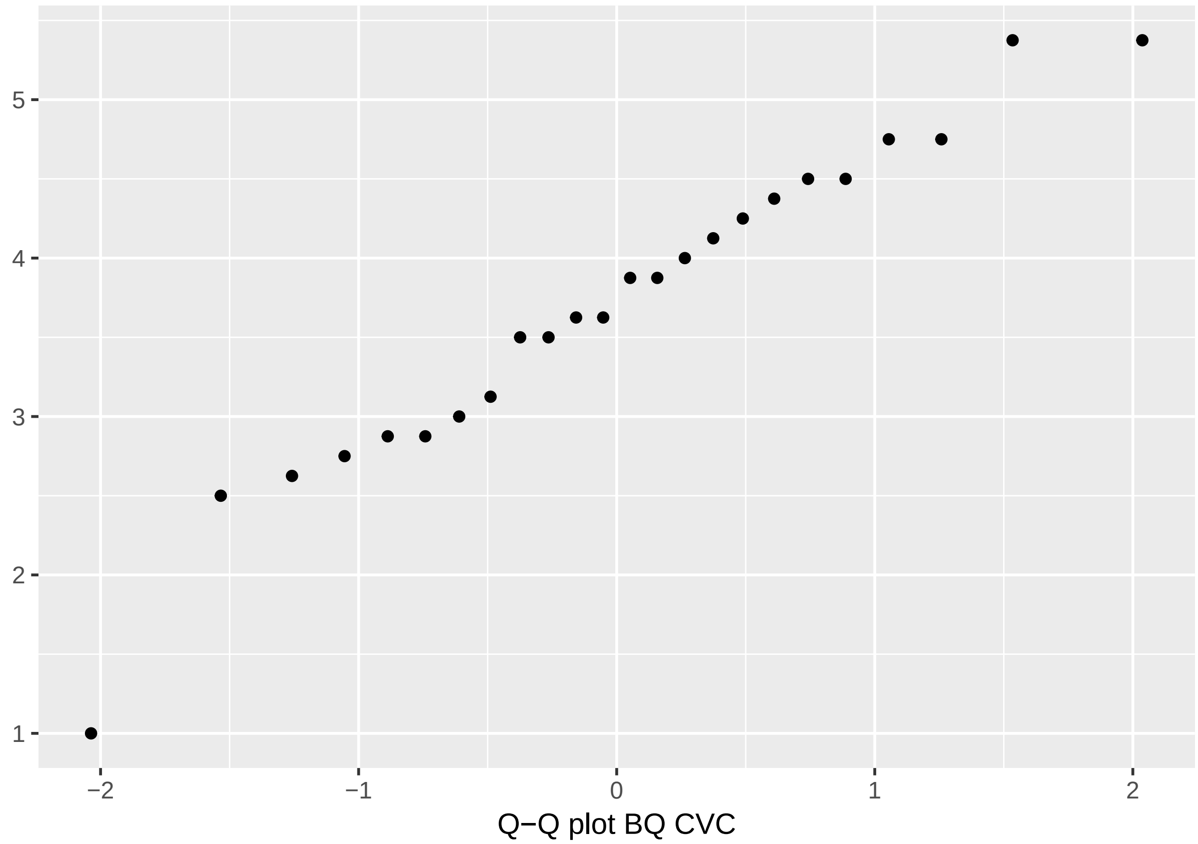



```r
ggplot(BQ_averages,
       aes(GAVC_average_score)) + geom_histogram(aes(y=..density..), binwidth = 1,
       colour="black", fill="white") + labs(x="Density plot BQ GAVC",
       y="Density") + stat_function(fun=dnorm,
       args=list(mean=mean(BQ_averages$GAVC_average_score, na.rm=TRUE),
       sd=sd(BQ_averages$GAVC_average_score, na.rm=TRUE)), colour="black", size=1)
```



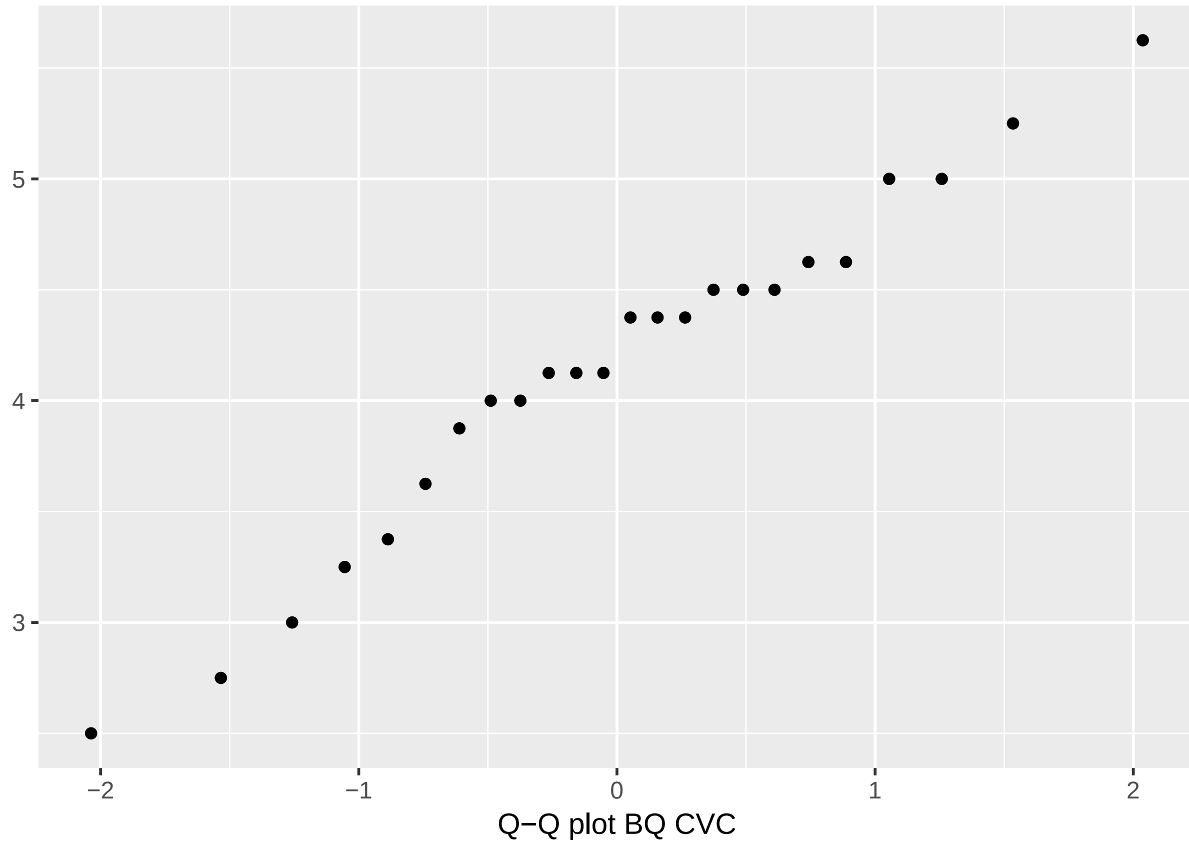

```
qplot(sample=BQ_averages$CVC_average_score, stat="qq") + labs(x="Q-Q plot BQ CVC")
```

```
## Warning: `stat` is deprecated
```




```
qplot(sample=BQ_averages$GAVC_average_score, stat="qq") + labs(x="Q-Q plot BQ CVC")
```

```
## Warning: `stat` is deprecated
```



By visually inspecting the histogram and qqplot it seems that the data is normally distributed. However, the sample size is rather small (n<30) so it is better to quantify the shape of the distribution:

```
round(stat.desc(BQ_averages[,c("CVC_average_score", "GAVC_average_score")],
                basic = FALSE, norm = TRUE), digits = 3)
```

```
##                CVC_average_score GAVC_average_score
## median                     3.750              4.250
## mean                       3.698              4.146
## SE.mean                    0.205              0.157
## CI.mean.0.95               0.425              0.325
## var                        1.011              0.592
## std.dev                    1.005              0.769
## coef.var                   0.272              0.186
## skewness                  -0.480             -0.348
## skew.2SE                  -0.508             -0.369
## kurtosis                   0.176             -0.451
## kurt.2SE                   0.096             -0.246
## normtest.W                 0.965              0.969
## normtest.p                 0.544              0.654
```

The skew.2SE and kurt.2SE are smaller than 0.98 (ignoring the plus or minus sign), which means the skew and kurtosis are not significant (at $p < 0.05$). The p-values (indicated by normtest.p) obtained by the Shapiro-Wilk test are >0.05 and thus we can assume that both the CVC and GAVC are normally distributed.

Paired difference test

The results were analysed using either a parametric or non-parametric paired difference test depending on the normality of the distribution. In this case the data is normally distributed and thus a paired t-test was used:

```
t.test(BQ_averages$CVC_average_score, BQ_averages$GAVC_average_score, paired = TRUE)
```

```
##
##  Paired t-test
##
## data:  BQ_averages$CVC_average_score and BQ_averages$GAVC_average_score
## t = -2.4325, df = 23, p-value = 0.02318
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.82883098 -0.06700236
## sample estimates:
## mean of the differences
##              -0.4479167
```

The p-value is 0.02 which is < 0.05 which means that we can reject the null hypothesis and the result is statistically significant.

Effect size

The effect size is calculated using Cohen's d:

```
cohensD(BQ_averages$CVC_average_score, BQ_averages$GAVC_average_score)
```

```
## [1] 0.5003649
```

The effect size (d) is 0.50.