

Blatt 8

Christian Peters

H17)

a)

Lese die Daten ein, standardisiere sie und berechne die Korrelationsmatrix.

```
data <- read.table('psych_results.txt', skip=14, header = TRUE)
data <- scale(data)
correlation_matrix <- cor(data)
```

Im Folgenden werden die verschiedenen Kriterien untersucht, um die Anzahl der Faktoren zu bestimmen.

Das erste Kriterium untersucht, wieviele der Eigenwerte der Korrelationsmatrix größer als 1 sind:

```
(eigen_values <- eigen(correlation_matrix)$values)
```

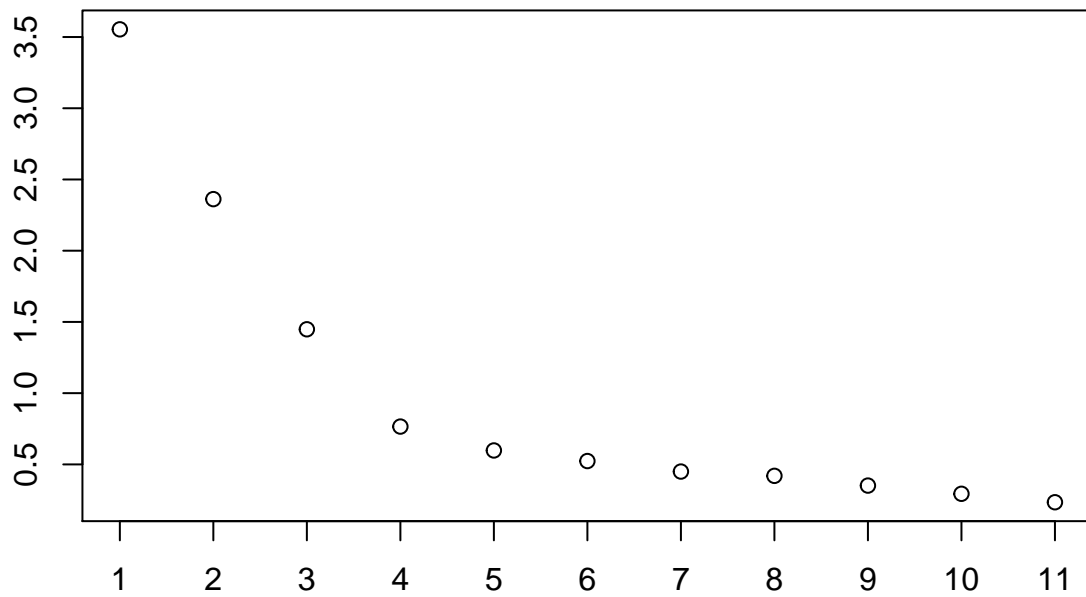
```
## [1] 3.5538455 2.3621133 1.4484420 0.7656200 0.5978845 0.5234879 0.4497023
## [8] 0.4196587 0.3513359 0.2935525 0.2343575
```

Hier würde man sich also für $p = 3$ Faktoren entscheiden.

Als nächstes kann der Scree-Graph betrachtet werden:

```
plot(eigen_values, main = 'Scree-Graph', xlab = '', ylab = '', xaxt = 'n')
axis(1, at = 1:length(eigen_values))
```

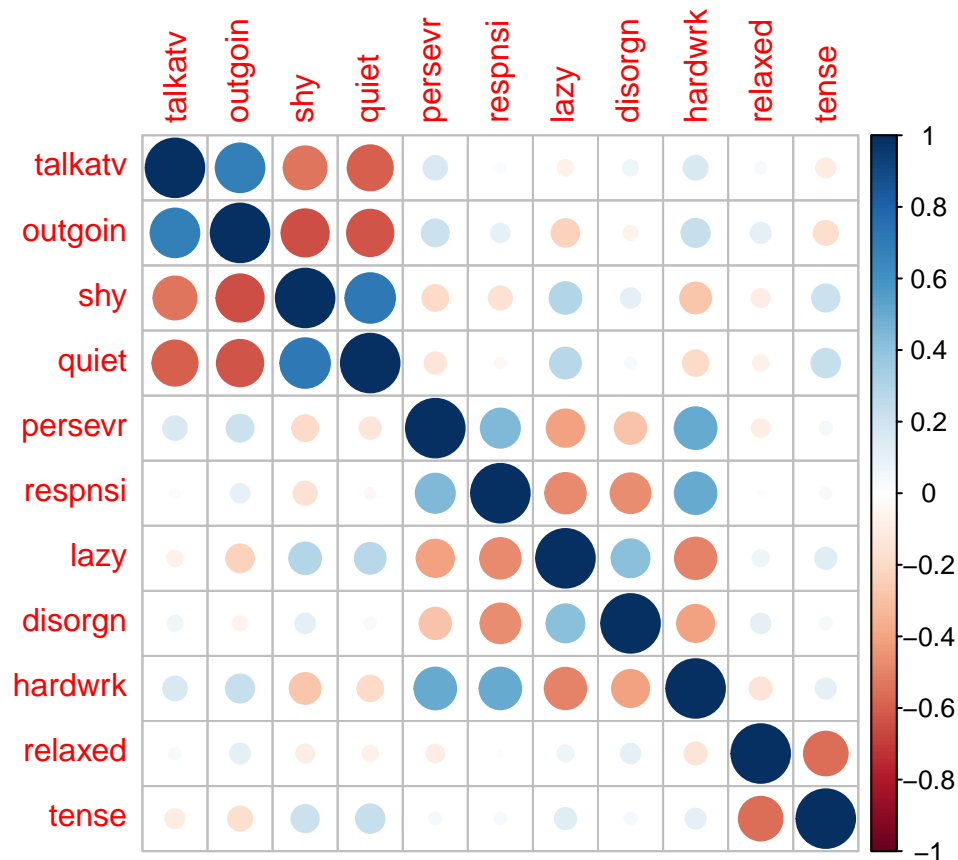
Scree-Graph



Auch hier ist es anhand des Knicks sinnvoll, sich für $p = 3$ Faktoren zu entscheiden.

Betrachte als letztes noch eine Visualisierung der Korrelationsmatrix:

```
corrplot(correlation_matrix)
```



Untersucht man hier, welche Merkmale gemeinsam korreliert sind, so lassen sich auch hier im Wesentlichen 3 Gruppen erkennen: talkatv und outgoing sind korreliert, quiet und shy bilden die zweite Gruppe und hardwrk, persevr und respnsi ließen sich ebenfalls zusammenfassen. Aufgrund dieser Beobachtungen ergibt es also auch hier Sinn, sich für ein Faktorenmodell mit $p = 3$ Faktoren zu entscheiden.

b)

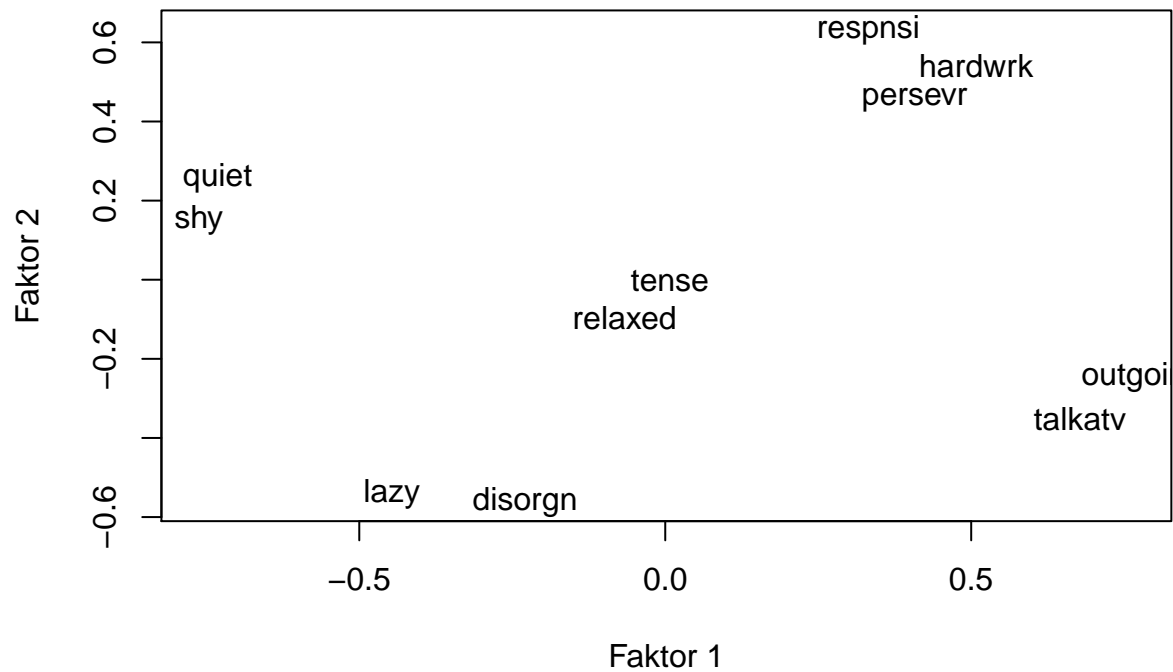
Passe nun ein Faktorenmodell mit 3 Faktoren ohne Rotation an:

```
factor_model <- factanal(data, 3, covmat=correlation_matrix, rotation = 'none')
```

Erstelle nun die Komponentendiagramme:

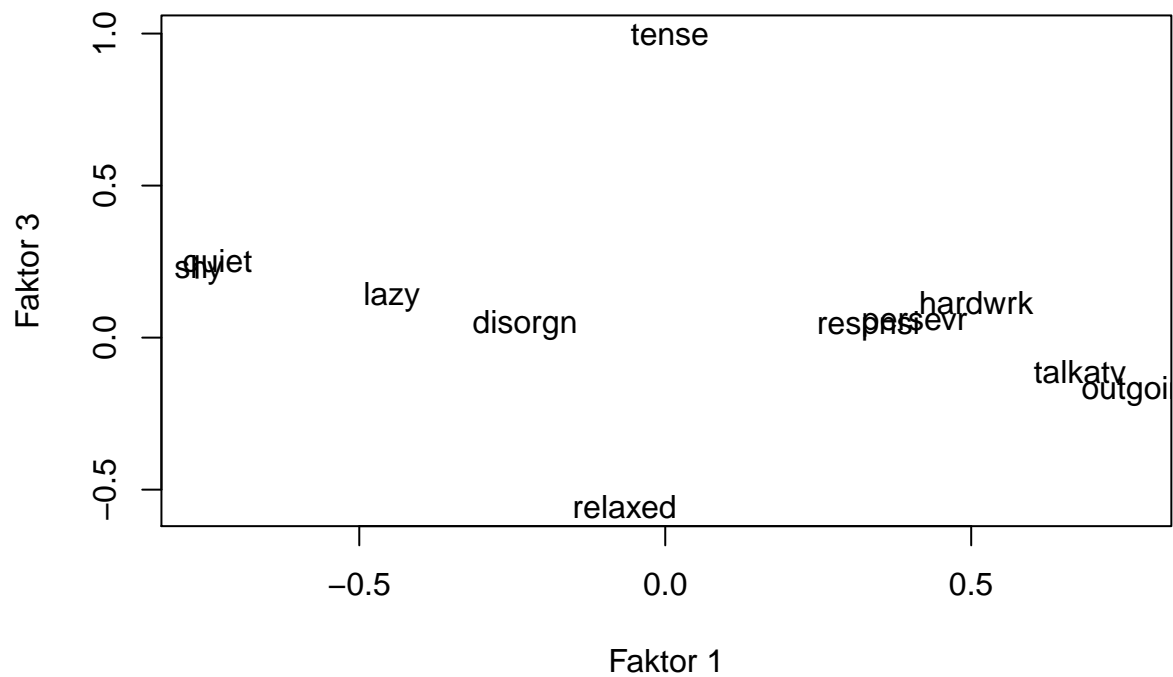
```
plot(factor_model$loadings[, 1], factor_model$loadings[, 2], type = 'n', xlab = 'Faktor 1',
      ylab = 'Faktor 2', main = 'Komponentendiagramm der Faktoren 1 und 2 ohne Rotation')
text(factor_model$loadings[, 1], factor_model$loadings[, 2], labels = colnames(data))
```

Komponentendiagramm der Faktoren 1 und 2 ohne Rotation



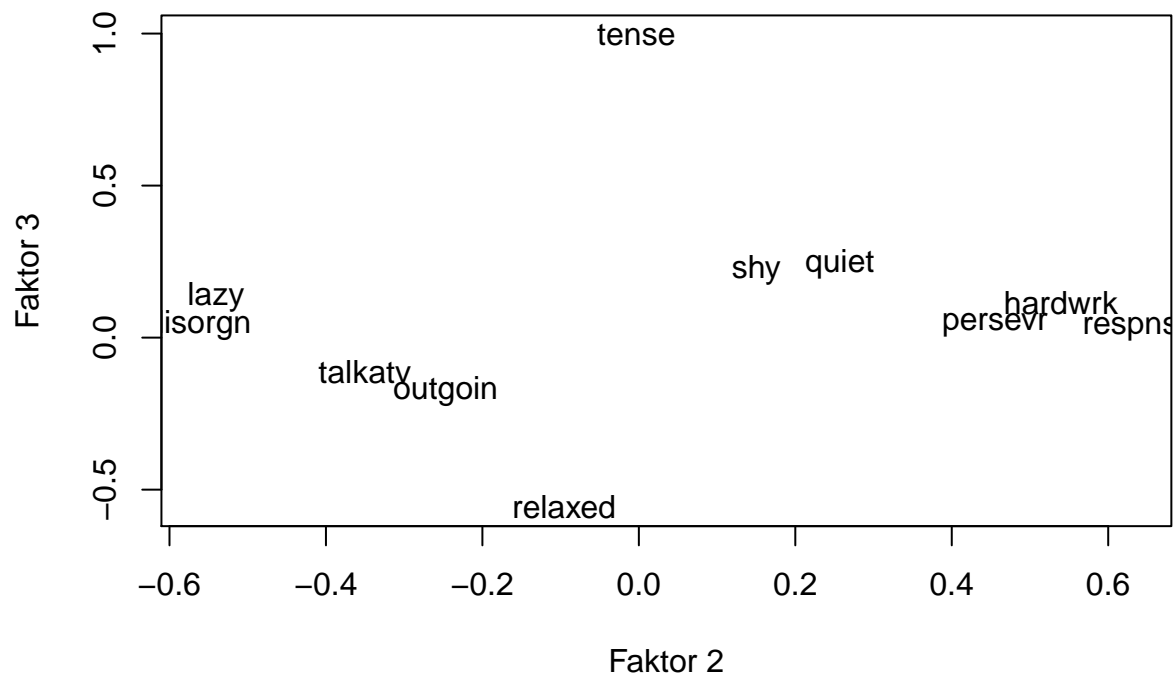
```
plot(factor_model$loadings[, 1], factor_model$loadings[, 3], type = 'n', xlab = 'Faktor 1',  
      ylab = 'Faktor 3', main = 'Komponentendiagramm der Faktoren 1 und 3 ohne Rotation')  
text(factor_model$loadings[, 1], factor_model$loadings[, 3], labels = colnames(data))
```

Komponentendiagramm der Faktoren 1 und 3 ohne Rotation



```
plot(factor_model$loadings[, 2], factor_model$loadings[, 3], type = 'n', xlab = 'Faktor 2',
      ylab = 'Faktor 3', main = 'Komponentendiagramm der Faktoren 2 und 3 ohne Rotation')
text(factor_model$loadings[, 2], factor_model$loadings[, 3], labels = colnames(data))
```

Komponentendiagramm der Faktoren 2 und 3 ohne Rotation



c)

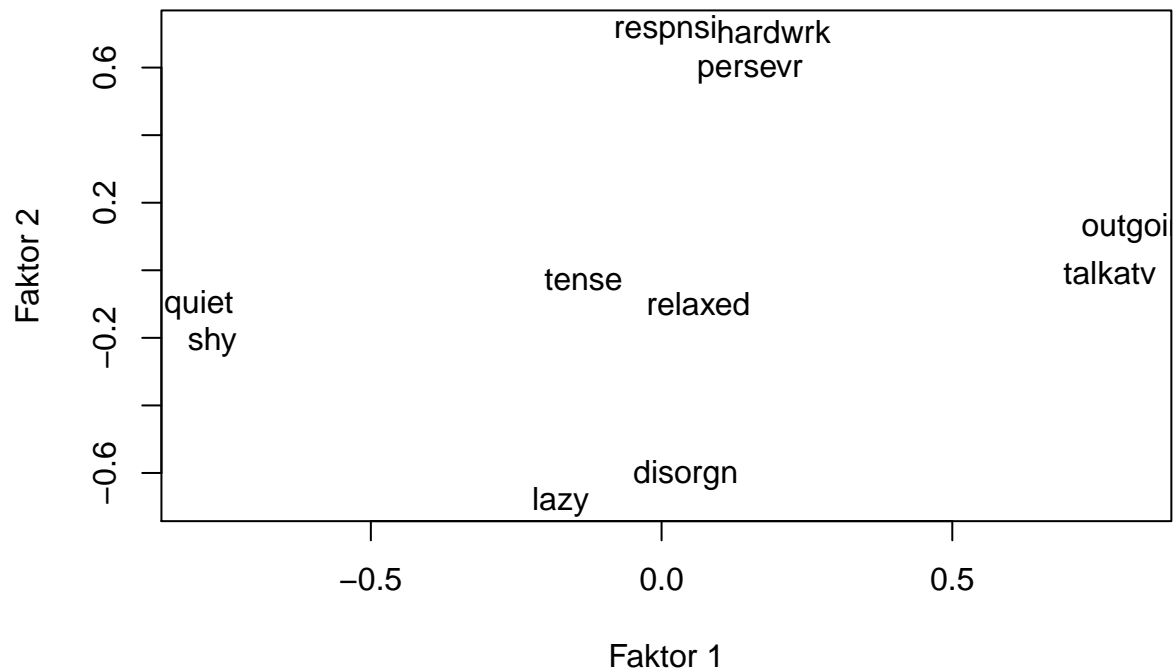
Verwende nun die Varimax-Rotation während der Anpassung:

```
factor_model <- factanal(data, 3, covmat=correlation_matrix, rotation = 'varimax')
```

Erstelle nun die Komponentendiagramme:

```
plot(factor_model$loadings[, 1], factor_model$loadings[, 2], type = 'n', xlab = 'Faktor 1',
      ylab = 'Faktor 2', main = 'Komponentendiagramm der Faktoren 1 und 2 mit Rotation')
text(factor_model$loadings[, 1], factor_model$loadings[, 2], labels = colnames(data))
```

Komponentendiagramm der Faktoren 1 und 2 mit Rotation



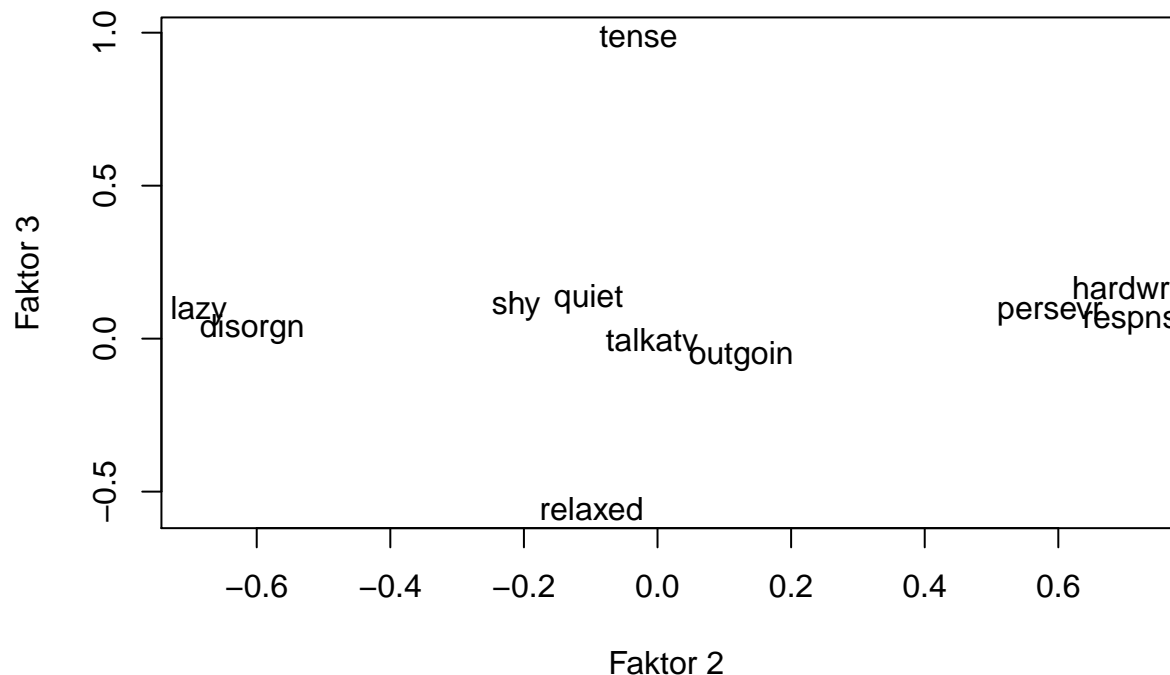
```
plot(factor_model$loadings[, 1], factor_model$loadings[, 3], type = 'n', xlab = 'Faktor 1',
      ylab = 'Faktor 3', main = 'Komponentendiagramm der Faktoren 1 und 3 mit Rotation')
text(factor_model$loadings[, 1], factor_model$loadings[, 3], labels = colnames(data))
```

Komponentendiagramm der Faktoren 1 und 3 mit Rotation



```
plot(factor_model$loadings[, 2], factor_model$loadings[, 3], type = 'n', xlab = 'Faktor 2',
      ylab = 'Faktor 3', main = 'Komponentendiagramm der Faktoren 2 und 3 mit Rotation')
text(factor_model$loadings[, 2], factor_model$loadings[, 3], labels = colnames(data))
```


Komponentendiagramm der Faktoren 2 und 3 mit Rotation



Man kann erkennen, dass Individuen stark positiv auf Faktor 1 laden, wenn sie große Werte bei den Attributen talkatv und outgoin haben. Individuen, bei denen die Attribute shy und quiet stärker ausgeprägt sind, laden stark negativ auf Faktor 1. Dies könnte bedeuten, dass Faktor 1 angibt, wie sehr extrovertiert eine Person ist. Große positive Werte von Faktor 1 deuten auf eine eher extrovertierte Persönlichkeit hin, große negative Werte hingegen eher auf eine introvertierte Persönlichkeit.

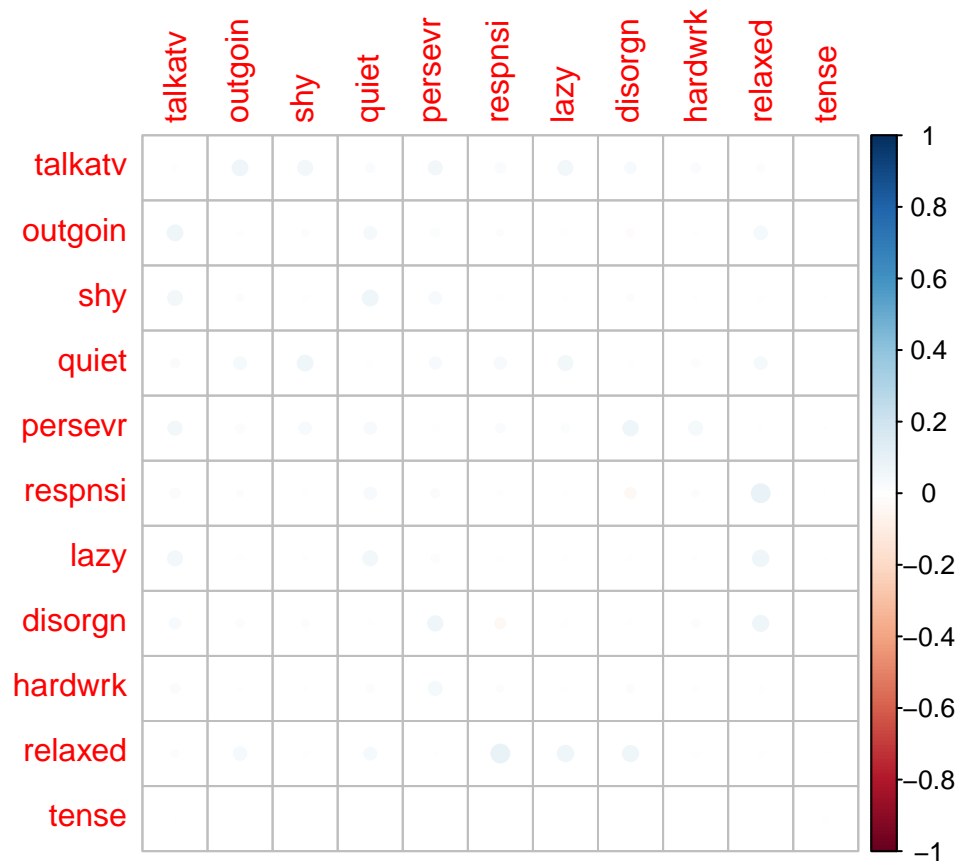
Faktor 2 scheint anzugeben, ob ein Individuum fleißig, verantwortungsbewusst und beharrlich ist. Individuen, bei denen diese Eigenschaften stark ausgeprägt sind, laden hoch auf diesen Faktor. Andersherum weisen Individuen, welche eher faul und unorganisiert sind, starke negative Ladungen auf diesen Faktor auf. Faktor 2 scheint also soetwas wie ein Maß für die Selbstdisziplin eines Individuums zu sein.

Für Faktor 3 sind nur zwei Merkmale relevant: tense und relaxed. Durch diesen Faktor werden also Individuen charakterisiert, die sich als besonders angespannt bezeichnen würden. Eher entspanntere Individuen laden im Gegenzug eher negativ auf diesen Faktor.

d)

Um zu beurteilen, wie gut die Faktoren die Korrelationsmatrix erklären, kann die Differenz zwischen der tatsächlichen und der modellierten Korrelationsmatrix $\hat{R} = LL' + V$ untersucht werden. Dies kann am besten durch eine Visualisierung dieser Differenz geschehen:

```
L <- factor_model$loadings
V <- diag(diag(correlation_matrix - L %*% t(L)))
difference <- correlation_matrix - (L %*% t(L) + V)
corrplot(difference)
```



Man erkennt also, dass die Abweichungen minimal sind. Alle relevanten Korrelationen werden gut durch das Modell erfasst. Das Modell scheint die Korrelationsmatrix also sehr gut zu erklären.