# Part3_Results

Meltem Ozcan

2024-09-12

This file assumes that the code in `Part1_Harmonization_continuous.Rmd` and `Part2_Harmonization_ordinal.Rmd` have been run. These two files result in the rds files that will be read in within the current file. If those files have not been run yet, please run the files to prepare the relevant rds files.

## Install and load packages, prepare data.

```
library(dplyr)
library(lavaan)
library(sirt)
library(mirt)
library(kableExtra)
library(here)
library(ggplot2)
# also install packages: haven, numDeriv
```

```
dat <- readRDS("rds/dat.rds")

est_partial <- readRDS("rds/est_partial.rds")
est_align <- readRDS("rds/est_align.rds")
est_partial_ord <- readRDS("rds/est_partial_ord.rds")
est_align_ord <- readRDS("rds/est_align_ord.rds")
est_align_ord_th <- readRDS("rds/est_align_ord_th.rds")
est_align_ord_lam <- readRDS("rds/est_align_ord_lam.rds")
fit_partial <- readRDS("rds/fit_partial.rds")
fit_align <- readRDS("rds/fit_align.rds")
mod_aligned <- readRDS("rds/mod_aligned.rds")

score_df_cont <- readRDS("rds/score_df_continuous.rds")
score_df_ord <- readRDS("rds/score_df_ord.rds")
df <- cbind(score_df_cont, score_df_ord[,c("approx_ord", "approx_ord_SE", "approx_ord_rel", "approx_ord_
head(df, 2)
```

```
##      stu_id sample sex dropout i1 i2 i3 i4 i5 mean_score partial_cont
## ELS1 101101    ELS   0       0  2  1  2  2  1        1.6    -1.114292
## ELS2 101102    ELS   0       0  4  3  4  4  4        3.8     1.658808
##      partial_SE approx_cont approx_SE partial_rel approx_rel partial_ev
## ELS1   0.278666   -1.131463 0.2781254    0.927941  0.9282002 0.07765475
## ELS2   0.278666    1.661305 0.2781254    0.927941  0.9282002 0.07765475
##       approx_ev approx_ord approx_ord_SE approx_ord_rel approx_ord_ev
## ELS1 0.07735375  -1.018915     0.2476258      0.9386815    0.05755857
## ELS2 0.07735375   1.312507     0.2987173      0.9107680    0.08126969
```

```r
saveRDS(df, "rds/df.rds")
m_items <- paste0("i", 1:5)
```

Also source the following R file to repeat the analyses up to this point on the dataset without any NAs for later sensitivity analyses.

```r
if (!all(file.exists("rds/score_df_noNAs.rds",
                     "rds/fit_partial_noNAs.rds",
                     "rds/fit_align_noNAs.rds",
                     "rds/est_align_ord_noNAs.rds"))) {
  source("code/repeatAnalysesNoNAs.R")
} else {
  complete_cases <- complete.cases(dat[, m_items]) & dat$sample == "ELS" |
    complete.cases(dat[, m_items[-3]]) & dat$sample == "HSLS"
  d_noNAs <- dat[complete_cases, ]
  fit_partial_noNAs <- readRDS("rds/fit_partial_noNAs.rds")
  fit_align_noNAs <- readRDS("rds/fit_align_noNAs.rds")
  est_align_ord_noNAs <- readRDS("rds/est_align_ord_noNAs.rds")

  score_df_noNAs <- readRDS("rds/score_df_noNAs.rds")
}
```

```
## Iteration: 1, Log-Lik: -143161.234, Max-Change: 2.93816Iteration: 2, Log-Lik: -116756.822, Max-Chang
## Iteration: 1, Log-Lik: -142906.991, Max-Change: 2.92118Iteration: 2, Log-Lik: -117479.479, Max-Chang
```

Table 1: Models by approach and dataset

| | | Continuous | | Ordinal | |
|---|---|---|---|---|---|
| | Mean Score | Partial | Approximate | Partial | Approximate |
| N = 30,749 | M_B1 | M1 | M2 | - | M3 |
| N = 29,202 | M_B2 | M4 | M5 | M6 | M7 |

```r
source('code/table_helper_functions.R')
```

Table 2: Loading estimates by approach and dataset

| | Continuous | | | | Ordinal | | | |
|---|---|---|---|---|---|---|---|---|
| | Partial (M1) | | Approximate (M2) | | Partial (M6) | | Approximate (M3) | |
| | ELS | HSLS | ELS | HSLS | ELS | HSLS | ELS | HSLS |
| i1 | 0.74 | 0.86 | 0.74 | 0.75 | 1.60 | 1.73 | 3.37 | 3.92 |
| i2 | 0.73 | 0.91 | 0.73 | 0.80 | 1.51 | 1.58 | 3.48 | 3.04 |
| i3 | 0.83 | NA | 0.83 | NA | 2.05 | NA | 4.37 | NA |
| i4 | 0.86 | 0.86 | 0.85 | 0.76 | 2.91 | 2.91 | 4.37 | 4.38 |
| i5 | 0.82 | 0.82 | 0.84 | 0.72 | 2.40 | 2.40 | 4.34 | 3.88 |

Table 3: Intercept estimates by approach and dataset

| | Continuous | |
|---|---|---|
| | Partial (M1) | Approximate (M2) |

| | ELS | HSLS | ELS | HSLS |
|---|---|---|---|---|
| i1 | 2.54 | 2.57 | 2.54 | 2.60 |
| i2 | 2.36 | 2.30 | 2.36 | 2.32 |
| i3 | 2.46 | NA | 2.46 | NA |
| i4 | 2.65 | 2.65 | 2.62 | 2.69 |
| i5 | 2.62 | 2.62 | 2.65 | 2.62 |

Table 6: Mean and SD of scores by approach and dataset

| | Mean Score | | Continuous | | | | Ordinal | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Partial | | Approximate | | Partial | | Approximate | |
| | M | SD | M | SD | M | SD | M | SD | M | SD |
| **N = 30,740** | **M_B1** | | **M1** | | **M2** | | | | **M3** | |
| ELS | 2.52 | 0.845 | -0.003 | 1.051 | -0.004 | 1.051 | - | - | -0.001 | 0.952 |
| HSLS | 2.938 | 0.661 | 0.469 | 0.763 | 0.501 | 0.869 | - | - | 0.619 | 0.908 |
| overall | 2.779 | 0.764 | 0.29 | 0.912 | 0.31 | 0.973 | - | - | 0.384 | 0.973 |
| **N = 29,202** | **M_B1** | | **M4** | | **M5** | | **M6** | | **M7** | |
| ELS | 2.545 | 0.833 | 0 | 1.038 | 0 | 1.038 | 0.001 | 0.939 | -0.001 | 0.957 |
| HSLS | 2.938 | 0.661 | 0.45 | 0.767 | 0.477 | 0.869 | 0.022 | 0.784 | 0.574 | 0.904 |
| overall | 2.797 | 0.751 | 0.289 | 0.9 | 0.306 | 0.96 | 0.015 | 0.843 | 0.368 | 0.964 |

Table 4: Threshold estimates by approach and dataset

| | Ordinal | | | |
|---|---|---|---|---|
| | Partial (M6) | | Approximate (M3) | |
| | ELS | HSLS | ELS | HSLS |
| i1\|t1 | -2.41 | -2.41 | 4.72 | 4.75 |
| i1\|t2 | 0.26 | -0.71 | 3.49 | 2.85 |
| i1\|t3 | 1.58 | 1.58 | 4.50 | NA |
| i2\|t1 | -1.74 | -1.74 | 5.64 | 5.63 |
| i2\|t2 | 0.45 | -0.05 | 5.74 | 4.94 |
| i2\|t3 | 1.90 | 1.90 | -0.44 | 0.35 |
| i3\|t1 | -2.33 | NA | -0.88 | -0.93 |
| i3\|t2 | 0.28 | NA | -0.48 | NA |
| i3\|t3 | 2.06 | NA | 0.29 | 1.24 |
| i4\|t1 | -4.01 | -4.01 | 0.44 | 0.56 |
| i4\|t2 | -0.16 | -1.68 | -3.07 | -5.61 |
| i4\|t3 | 2.39 | 2.39 | -4.18 | -5.50 |
| i5\|t1 | -3.39 | -3.39 | -4.26 | NA |
| i5\|t2 | -0.23 | -1.10 | -3.56 | -5.78 |
| i5\|t3 | 1.92 | 2.32 | -3.41 | -5.78 |

Table 5: Latent mean and variance estimates by approach and dataset

| | Continuous | | | | Ordinal | | | |
|---|---|---|---|---|---|---|---|---|
| | Partial (M1) | | Approximate (M2) | | Partial (M6) | | Approximate (M3) | |
| | ELS | HSLS | ELS | HSLS | ELS | HSLS | ELS | HSLS |
| Latent mean | 0 | 0.47 | 0 | 0.50 | 0 | 0.28 | 0 | 0.70 |
| Latent variance | 1 | 0.53 | 1 | 0.69 | 1 | 0.66 | 1 | 0.94 |

Table 7: Correlation of FS (N = 30,740)

| | | Continuous | | Ordinal | |
|---|---|---|---|---|---|
| | Mean scores | Partial | Approximate | Partial | Approximate |
| | **M_B1** | **M1** | **M2** | **-** | **M3** |
| **M_B1** | 1 | 0.992 | 0.993 | - | 0.989 |
| **M1** | 0.992 | 1 | 0.998 | - | 0.98 |
| **M2** | 0.993 | 0.998 | 1 | - | 0.986 |
| **-** | - | - | - | - | - |
| **M3** | 0.989 | 0.98 | 0.986 | - | 1 |

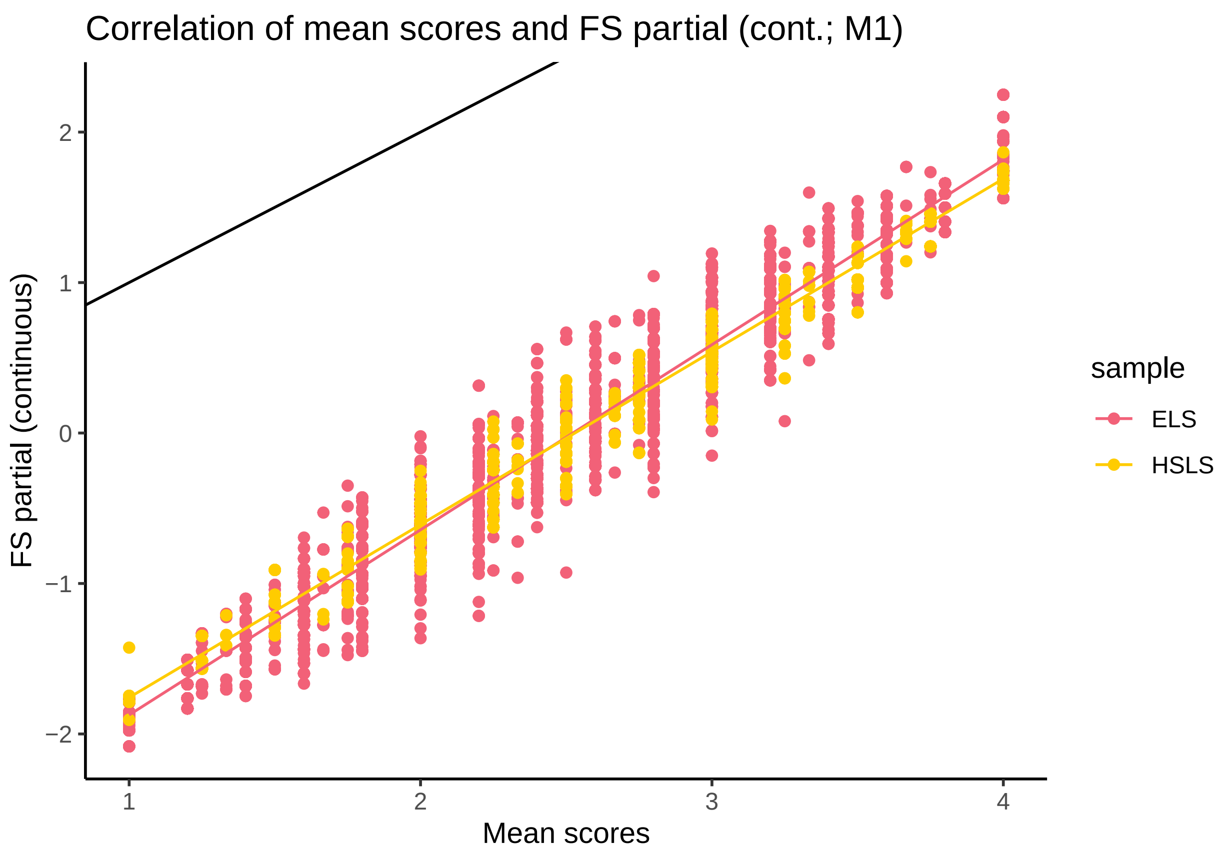
Table 8: Correlation of FS on dataset with no NAs (N = 29,202)

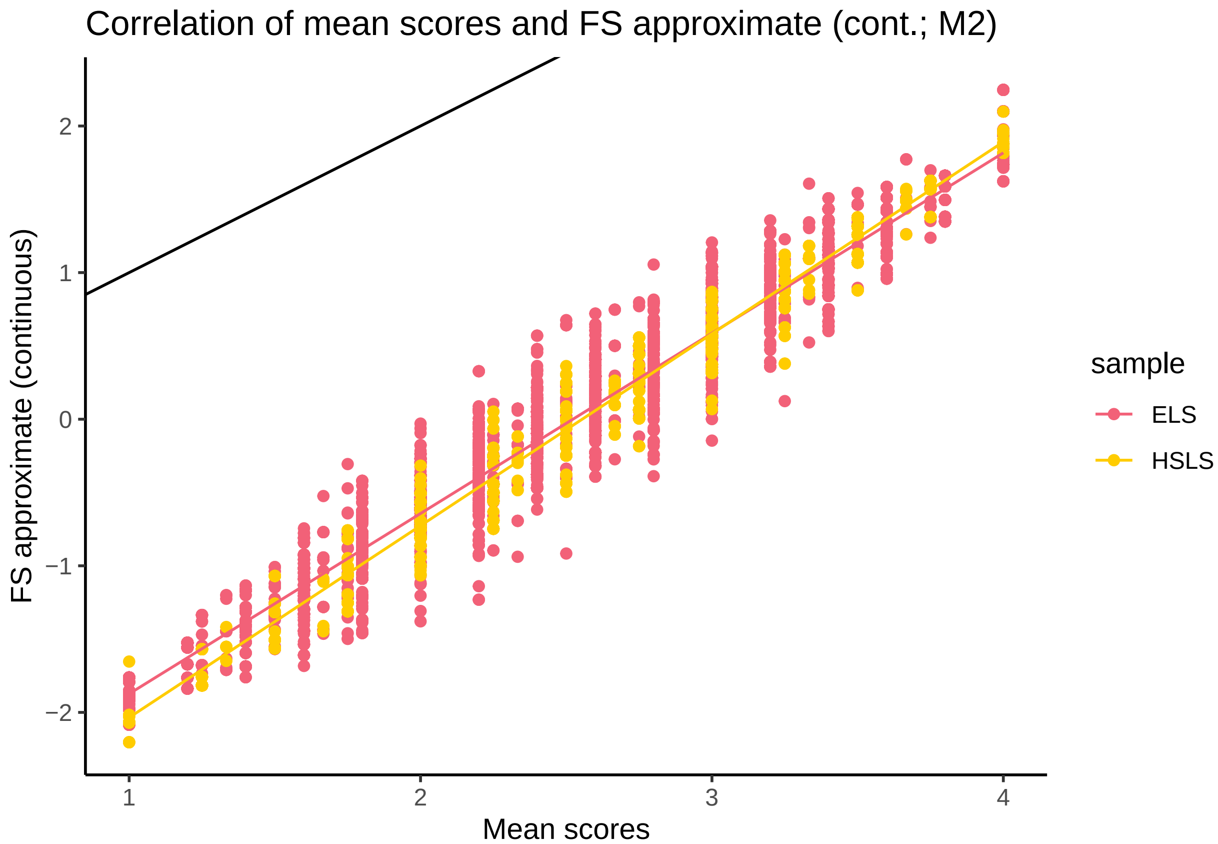
| | | Continuous | | Ordinal | |
|---|---|---|---|---|---|
| | Mean scores | Partial | Approximate | Partial | Approximate |
| | **M_B2** | **M4** | **M5** | **M6** | **M7** |
| **M_B2** | 1 | 0.992 | 0.993 | 0.965 | 0.991 |
| **M4** | 0.992 | 1 | 0.998 | 0.964 | 0.983 |
| **M5** | 0.993 | 0.998 | 1 | 0.966 | 0.988 |
| **M6** | 0.965 | 0.964 | 0.966 | 1 | 0.957 |
| **M7** | 0.991 | 0.983 | 0.988 | 0.957 | 1 |

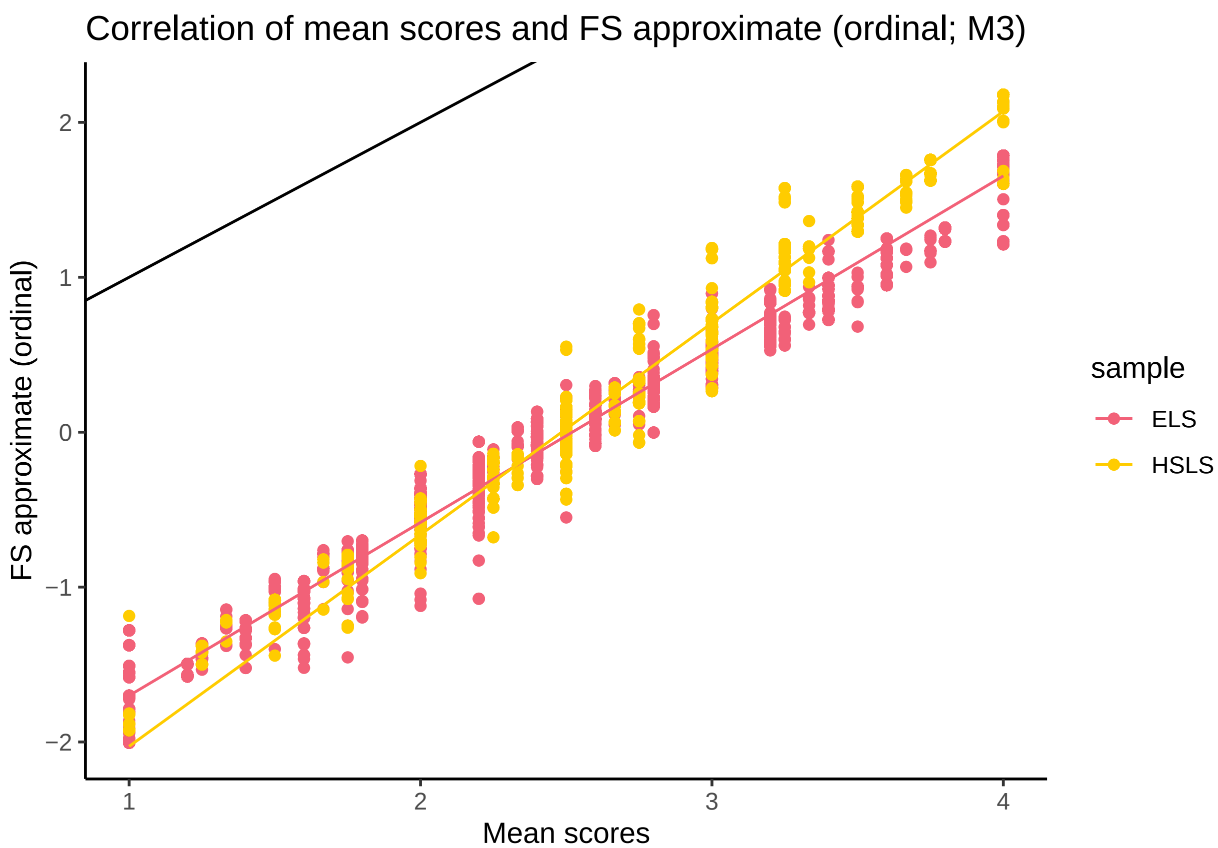
## Score distributions

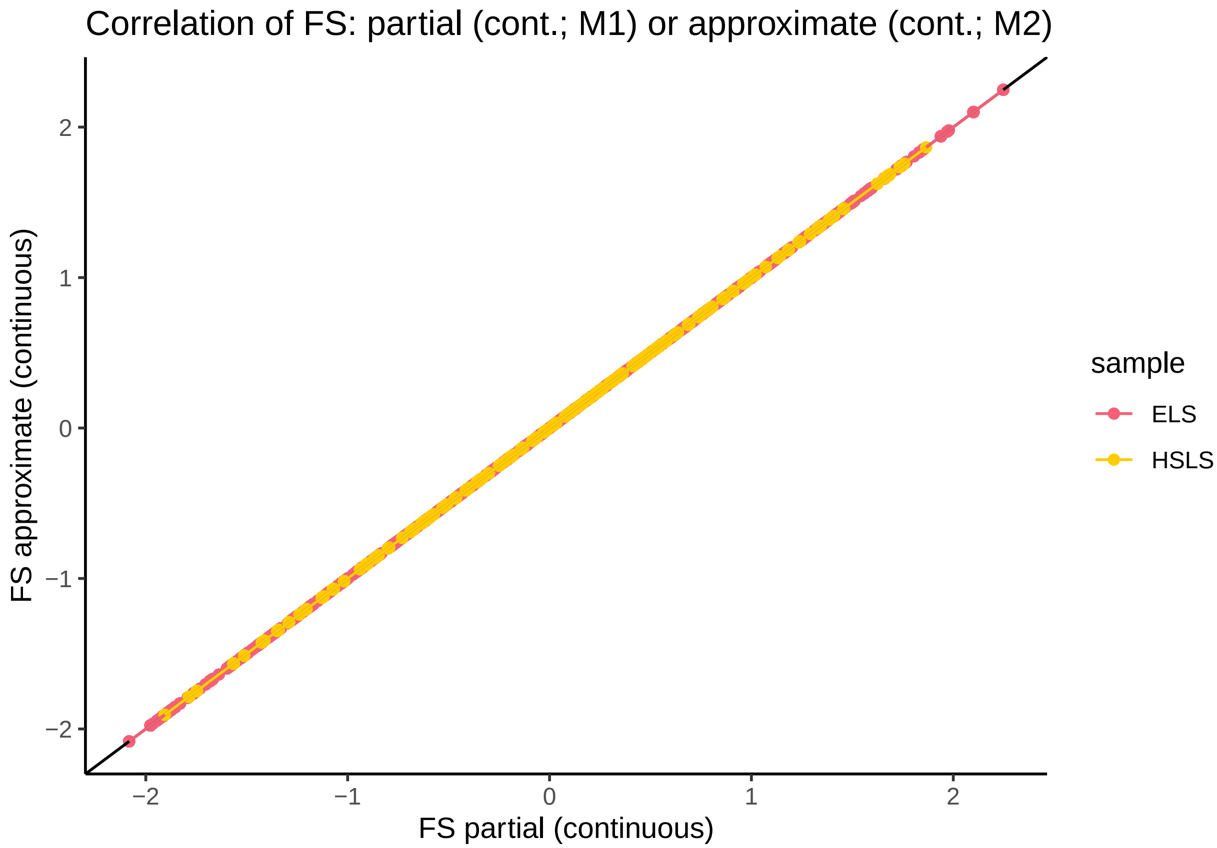
## Score correlations

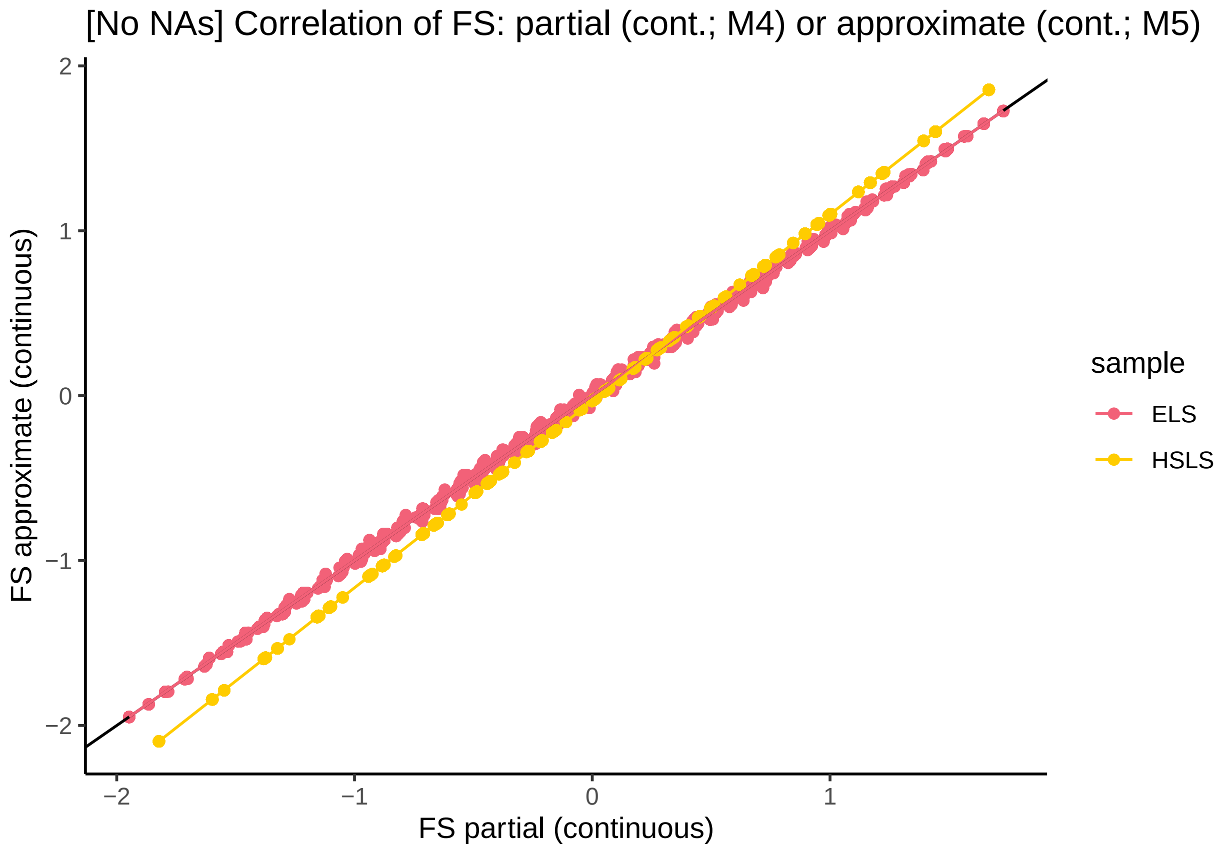
```
# compute matrix of correlations
cor_30749 <- cor(df[, c("mean_score", "partial_cont", "approx_cont", "approx_ord")], use = "complete.obs
```


Correlation of mean scores and FS partial (cont.; M1)
FS partial (continuous)
Mean scores
2
1
0
−1
−2
1
2
3
4
sample
ELS
HSLS

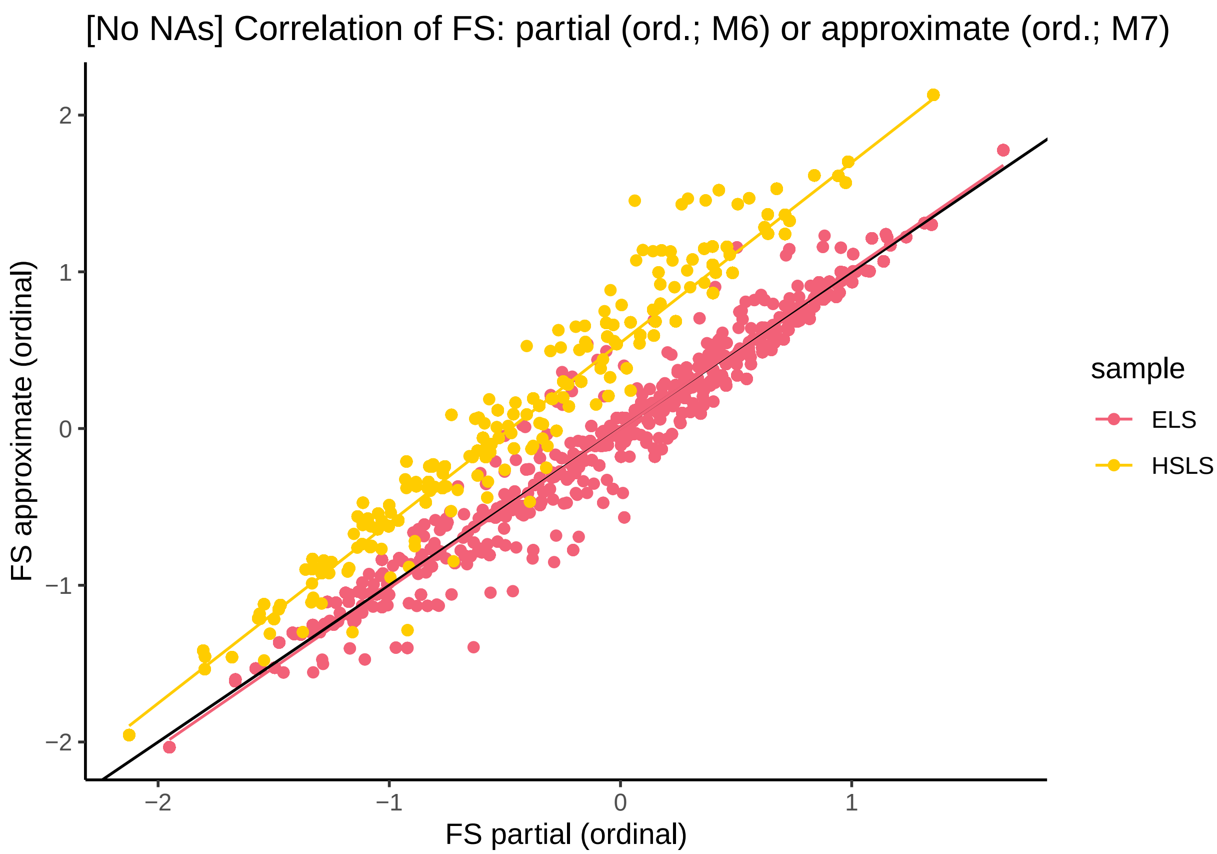

Correlation of mean scores and FS approximate (cont.; M2)
FS approximate (continuous)
Mean scores
2
1
0
−1
−2
1
2
3
4
sample
ELS
HSLS

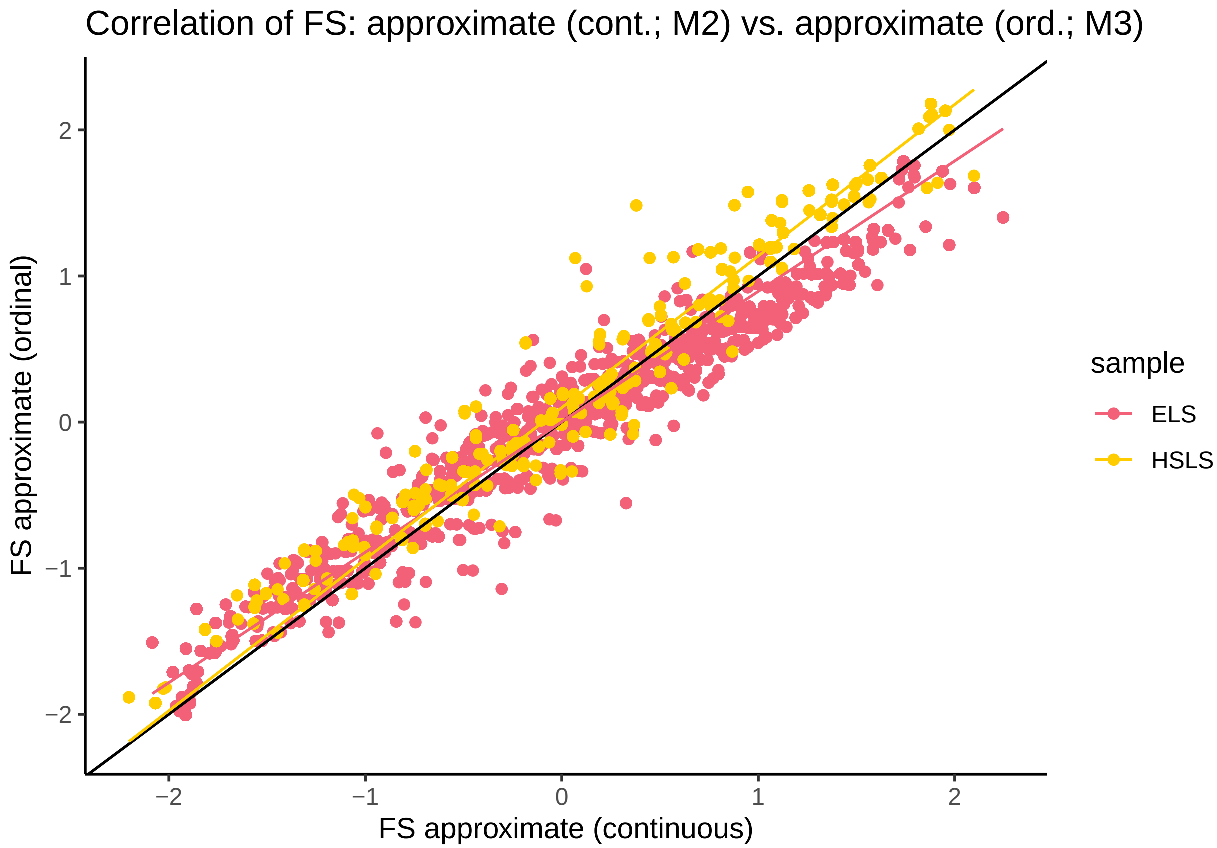

Correlation of mean scores and FS approximate (ordinal; M3)
FS approximate (ordinal)
Mean scores
sample
ELS
HSLS
2
1
0
−1
−2
1
2
3
4


Correlation of FS: partial (cont.; M1) or approximate (cont.; M2)
FS approximate (continuous)
FS partial (continuous)
2
1
0
−1
−2
−2
−1
0
1
2
sample
ELS
HSLS



[No NAs] Correlation of FS: partial (cont.; M4) or approximate (cont.; M5)
FS approximate (continuous)
FS partial (continuous)
2
1
0
−1
−2
−2
−1
0
1
sample
ELS
HSLS


[No NAs] Correlation of FS: partial (ord.; M6) or approximate (ord.; M7)
FS approximate (ordinal)
FS partial (ordinal)
2
1
0
−1
−2
−2
−1
0
1
sample
ELS
HSLS


Correlation of FS: approximate (cont.; M2) vs. approximate (ord.; M3)
FS approximate (ordinal)
FS approximate (continuous)
−2
−1
0
1
2
sample
ELS
HSLS

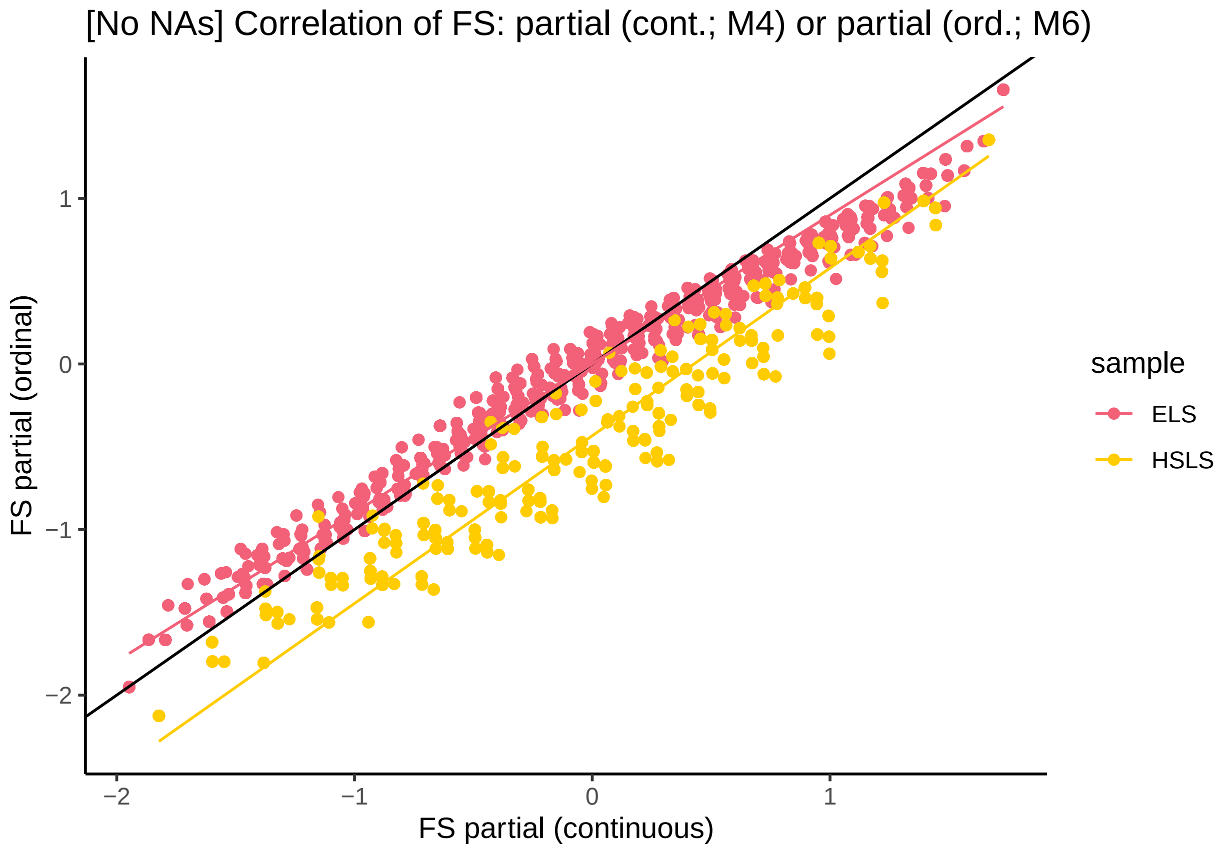
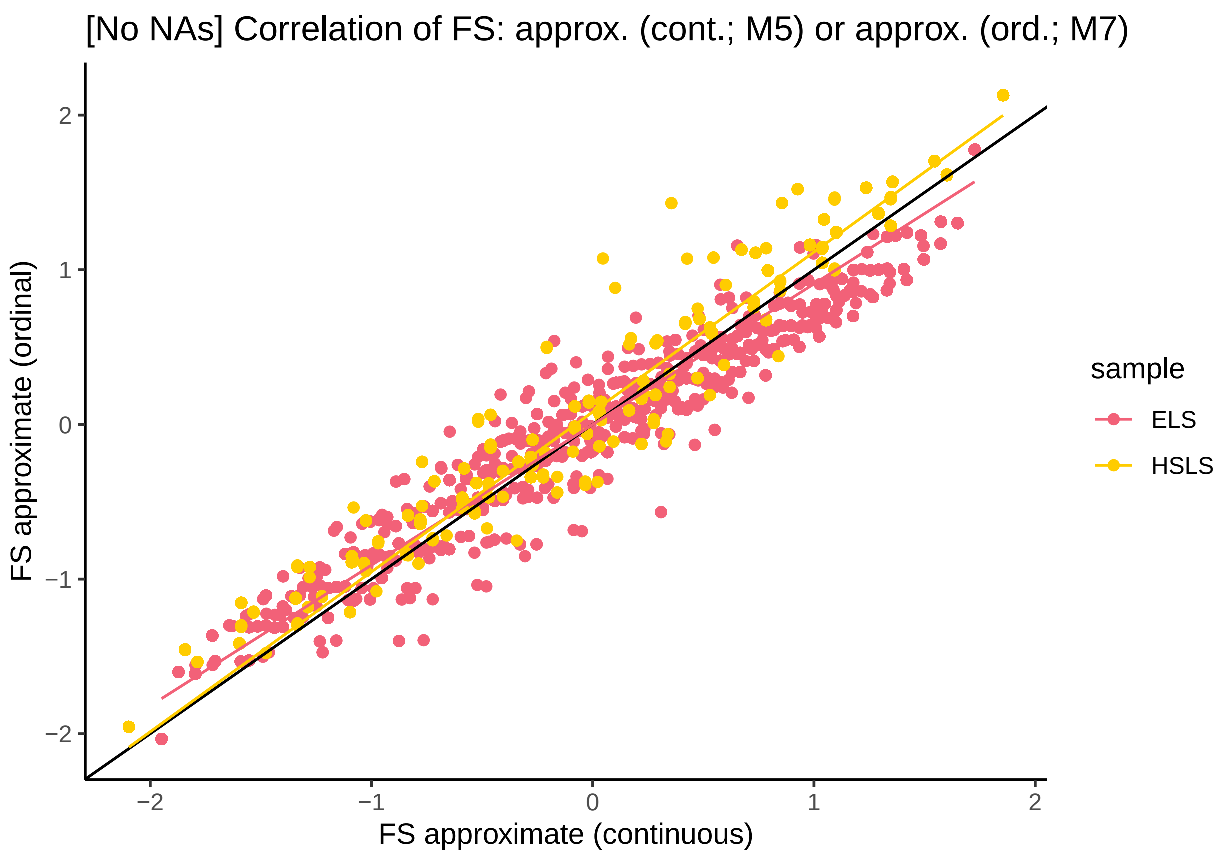

[No NAs] Correlation of FS: partial (cont.; M4) or partial (ord.; M6)
FS partial (ordinal)
FS partial (continuous)
1
0
−1
−2
−2
−1
0
1
sample
ELS
HSLS

Table 9: Mean and SD of FS Reliabilities (N = 30,740)

| | Continuous | | | | Ordinal | | | |
|---|---|---|---|---|---|---|---|---|
| | Partial | | Approximate | | Partial | | Approximate | |
| | ELS | HSLS | ELS | HSLS | ELS | HSLS | ELS | HSLS |
| Mean | 0.9139 | 0.9195 | 0.914 | 0.9193 | - | - | 0.9068 | 0.8833 |
| SD | 0.0515 | 0.0102 | 0.0519 | 0.0102 | - | - | 0.0668 | 0.0418 |




## Reliability

```
mnsd(df$partial_rel)
```

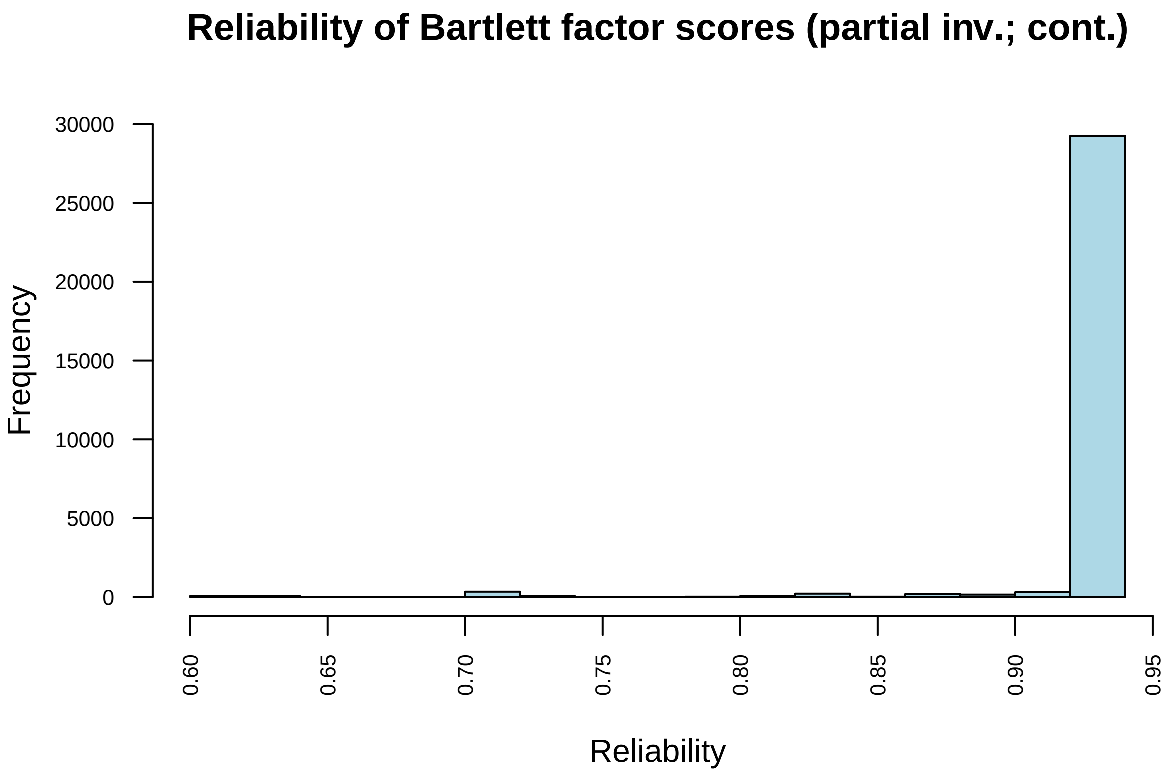
```
## [1] 0.91736732 0.03280564
```

```
mnsd(df$approx_rel)
```

```
## [1] 0.91732712 0.03304773
```

```
mnsd(df$approx_ord_rel)
```

```
## [1] 0.89223975 0.05391562
```


Reliability of Bartlett factor scores (partial inv.; cont.)
Frequency
30000
25000
20000
15000
10000
5000
0
0.60
0.65
0.70
0.75
0.80
0.85
0.90
0.95
Reliability

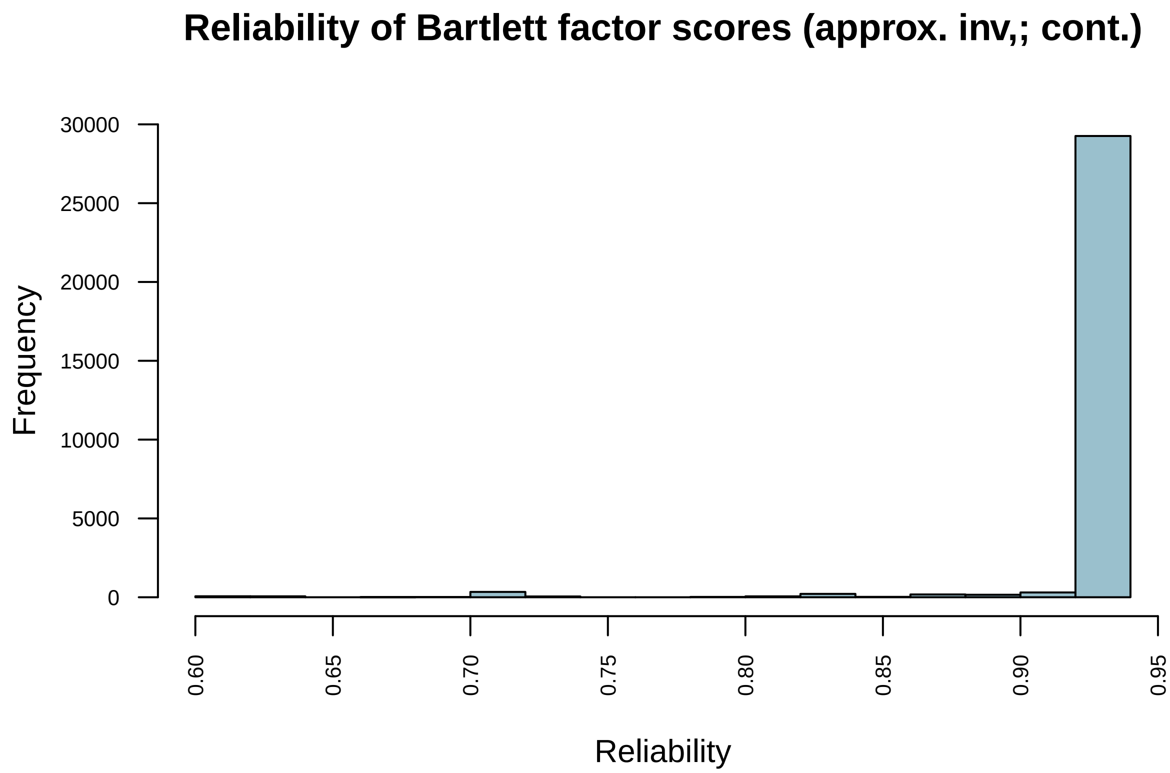
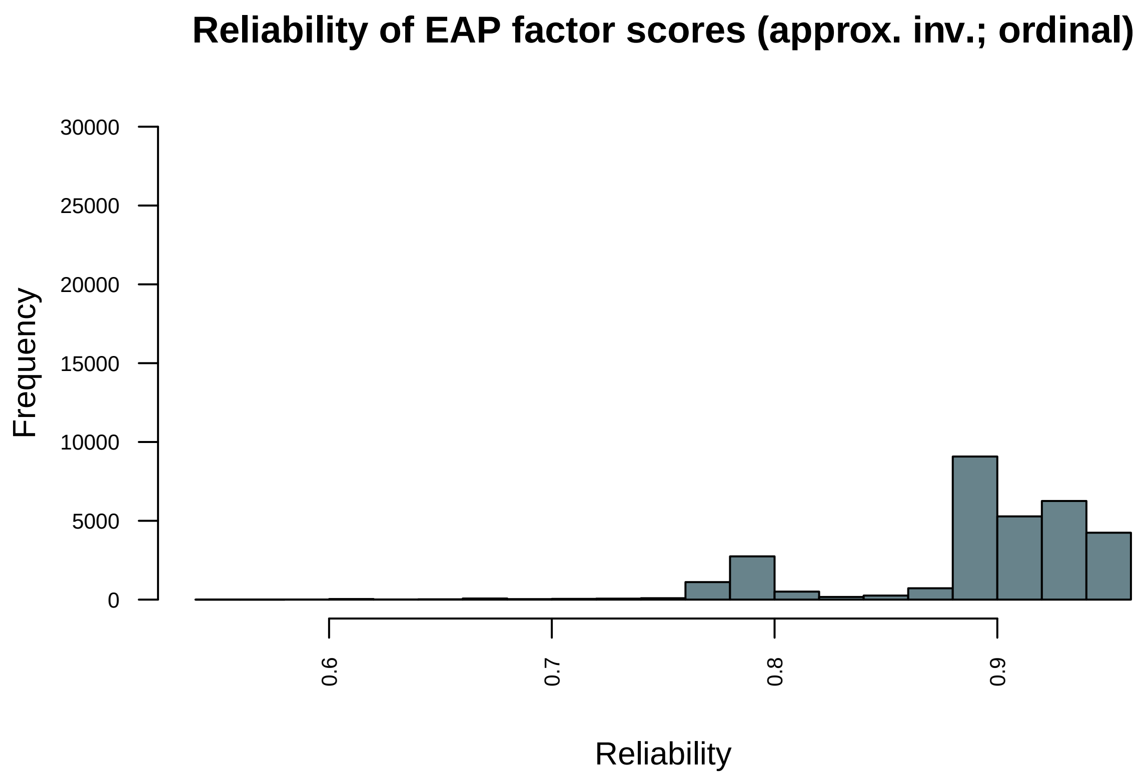
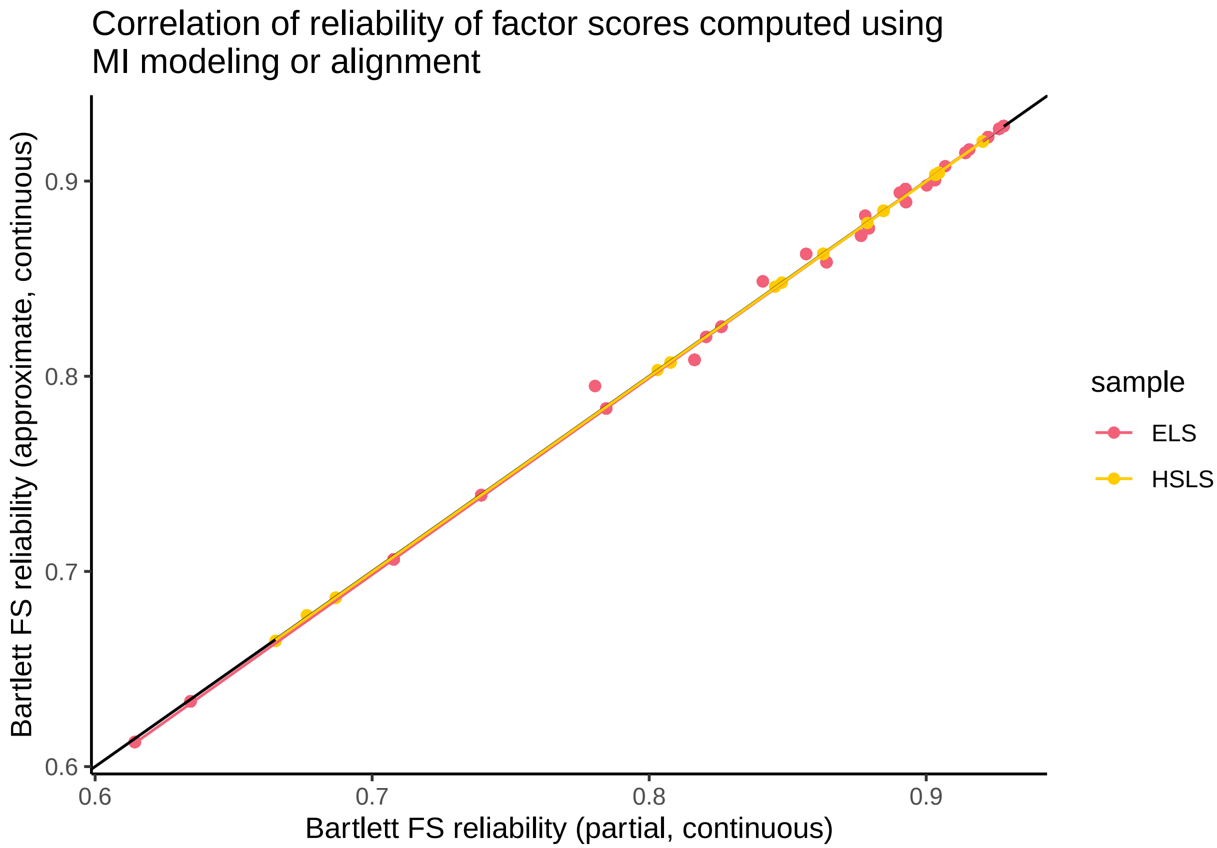

Reliability of Bartlett factor scores (approx. inv,; cont.)
Frequency
30000
25000
20000
15000
10000
5000
0
0.60
0.65
0.70
0.75
0.80
0.85
0.90
0.95
Reliability

Table 10: Correlation of FS Reliabilities (N = 30,740)

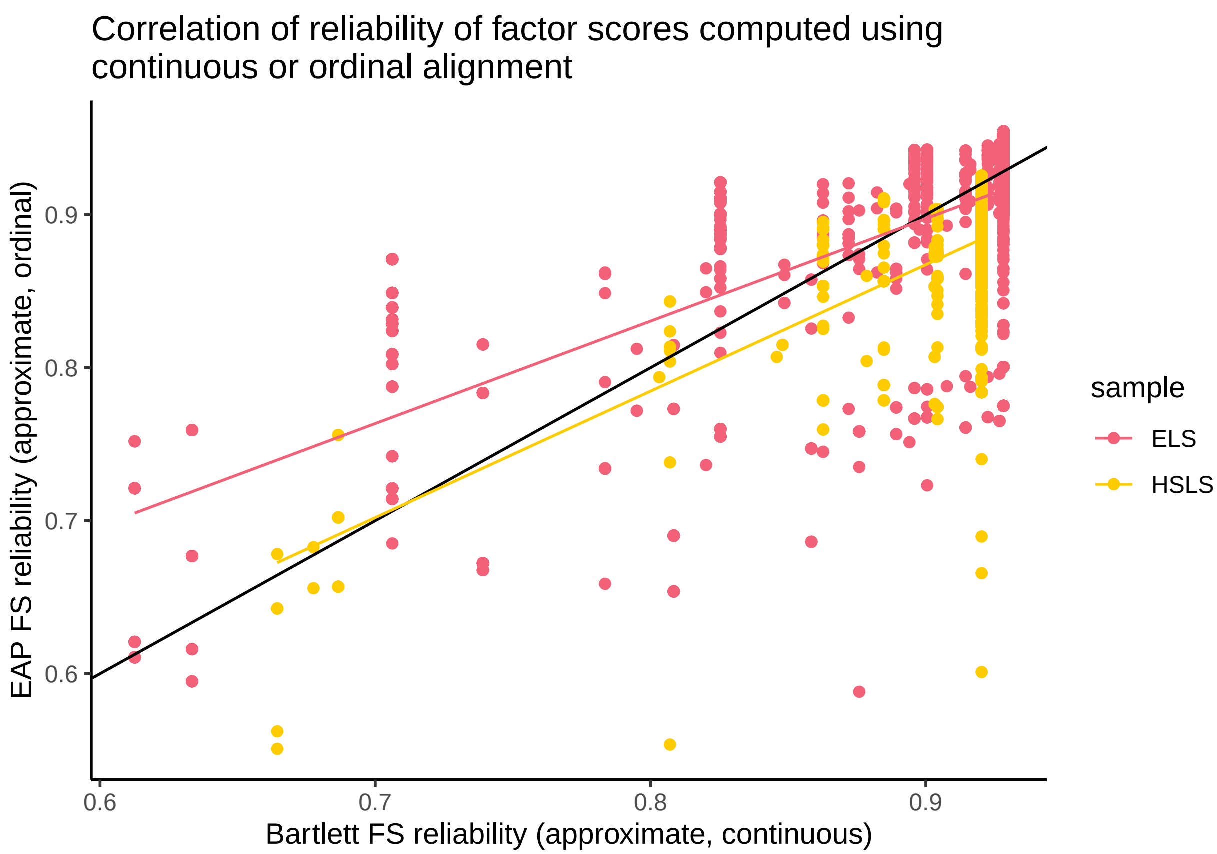
| | Continuous | | Ordinal | |
|---|---|---|---|---|
| | Partial | Approximate | Partial | Approximate |
| | **M1** | **M2** | **-** | **M3** |
| **M1** | 1 | 0.9999 | - | 0.3952 |
| **M2** | 0.9999 | 1 | - | 0.3971 |
| **-** | - | - | - | - |
| **M3** | 0.3952 | 0.3971 | - | 1 |




```
rel_30749 <- cor(as.matrix(apply(df[ ,c("partial_rel",
                                        "approx_rel",
                                        "approx_ord_rel")],
                                  as.numeric, MARGIN = 2)),
                 use = "complete.obs")
rel_30749 <- round(rel_30749, 4)
rel_30749_add_row <- rbind(rel_30749[1:2,], rep("-",3) , rel_30749[3,])
rel_30749_add_col <- cbind(rel_30749_add_row[,1:2], rep("-",4) , rel_30749_add_row[,3])
colnames(rel_30749_add_col) <- rownames(rel_30749_add_col) <- as.character(models[1,])[2:5]
```



Correlation of reliability of factor scores computed using
MI modeling or alignment
Bartlett FS reliability (approximate, continuous)
0.9
0.8
0.7
0.6
0.6
0.7
0.8
0.9
Bartlett FS reliability (partial, continuous)
sample
ELS
HSLS




### Dmacs effect sizes of noninvariance

We use the R package `pinsearch` to compute the Dmacs effect size of noninvariance for each approach. We use the `pin_effsize()` function and `dmacs_ordered()` for the ordinal case. Note that if there are more than two groups, `pin_effsize()` computes the Fmacs effect size when data are continuous, and `fmacs_ordered()` should replace `dmacs_ordered()` in the ordinal case.

#### Continuous case

```
dmacs_partial <- pinsearch::pin_effsize(fit_partial)
dmacs_align <- pinsearch::pin_effsize(fit_align)
```

#### Ordinal case

```
# indicate which thresholds belong to which item
est_align_ord_th_sub <- est_align_ord_th[, -c(3, 8, 13)]
colnames(est_align_ord_th_sub) <- rep(c(1, 2, 4, 5), 3)

# compute pooled SD for dmacs_ordered()
vars <- apply(dat[m_items], MARGIN = 2,
              FUN = \(x) tapply(x, INDEX = dat$sample, FUN = var, na.rm = TRUE)
              )
(wgt_mat <- as.matrix(rbind(
  colSums(!is.na(dat[dat$sample=="ELS", m_items])),
  colSums(!is.na(dat[dat$sample=="HSLS", m_items]))
```

```r
)))
```

```
##          i1    i2    i3    i4    i5
## [1,] 11391 11432 11047 10823 10661
## [2,] 19049 19006     0 18926 18976
```

```r
item_n <- ifelse(wgt_mat > 0, wgt_mat, NA)
pooled_sd <- sqrt(colSums(vars * (item_n - 1), na.rm = TRUE) /
                    colSums(item_n - 1, na.rm = TRUE)
                  )

dmacs_align_ord <- pinsearch::dmacs_ordered(
  thresholds = as.matrix(est_align_ord_th_sub),
  loadings = as.matrix(est_align_ord_lam)[, -3],
  link = "logit",
  pooled_item_sd = pooled_sd[-3]
)
```

```r
# compute dmacs effect size for ordinal MI modeling approach:
# first recompute pooled SD since rows with NAs had to be dropped in the ordinal
# MI modeling approach
## note that now we're using d_noNAs as out dataset.



vars_noNAs <- apply(d_noNAs[m_items], MARGIN = 2,
                    FUN = \(x) tapply(x, INDEX = d_noNAs$sample,
                                      FUN = var, na.rm = TRUE)
              )
(wgt_mat_noNAs <- as.matrix(rbind(
  colSums(!is.na(d_noNAs[d_noNAs$sample=="ELS", m_items])),
  colSums(!is.na(d_noNAs[d_noNAs$sample=="HSLS", m_items]))
)))
```

```
##          i1    i2    i3    i4    i5
## [1,] 10443 10443 10443 10443 10443
## [2,] 18759 18759     0 18759 18759
```

```r
item_n_noNAs <- ifelse(wgt_mat_noNAs > 0, wgt_mat_noNAs, NA)
pooled_sd_noNAs <- sqrt(colSums(vars_noNAs * (item_n_noNAs - 1), na.rm = TRUE) /
                          colSums(item_n_noNAs - 1, na.rm = TRUE)
                        )
# indicate which thresholds belong to which item
est_partial_ord_th_sub <- as.matrix(rbind(t(est_partial_ord$ELS$tau)[-c(7:9)],
                                          t(est_partial_ord$HSLS$tau)))
colnames(est_partial_ord_th_sub) <- c(rep(1, 3), rep(2, 3), rep(4, 3), rep(5, 3))
est_partial_ord_lam <- as.matrix(rbind(t(est_partial_ord$ELS$lambda)[-3],
                                       t(est_partial_ord$HSLS$lambda)))

dmacs_partial_ord <- pinsearch::dmacs_ordered(
  thresholds = est_partial_ord_th_sub,
  loadings = est_partial_ord_lam,
  link = "logit",
  pooled_item_sd = pooled_sd_noNAs[-3]
)
```

Table 11: Dmacs effect sizes of noninvariance by approach (N = 30,749)

| | Continuous | | Ordinal | |
|---|---|---|---|---|
| | Partial (M1) | Approximate (M2) | - | Approximate (M3) |
| i1 | 0.149 | 0.072 | - | 0.217 |
| i2 | 0.219 | 0.09 | - | 0.206 |
| i4 | - | 0.137 | - | 0.267 |
| i5 | - | 0.15 | - | 0.311 |

Table 12: Dmacs effect sizes of noninvariance by approach (N = 29,202)

| | Continuous | | Ordinal | |
|---|---|---|---|---|
| | Partial (M4) | Approximate (M5) | Partial (M6) | Approximate (M7) |
| i1 | 0.147 | 0.079 | 0.212 | 0.208 |
| i2 | 0.215 | 0.091 | 0.113 | 0.187 |
| i4 | - | 0.133 | 0.263 | 0.263 |
| i5 | - | 0.145 | 0.142 | 0.298 |