APPROACHING THE MAIN EXAM

1. Split your time according to the marks for the sections and questions.

2. If you get stuck on an MCQ, leave it and come back to it. It is much easier at the end of the exam to return quickly to an MCQ than it is to return quickly to providing some interpretation of results.

3. Typically, students pick up more marks in the first parts of an answer than in the latter sections (we often repeat ourselves!). Make an attempt to answer all questions, even if on the initial attempt, your answer is brief.

4. Read the questions in Section B carefully. For the longer questions, the question will include information on the things we expect your answers to comment on. Questions in Section B have the number of marks indicated next to each question. As a (very rough) rule of thumb, a mark will typically correspond to a sentence (so for a 4-mark question a 4-5 sentence answer would likely be appropriate).

5. You will not necessarily need to directly use all the equations on the equation sheet, but they may help you answer questions.

Section A: Multiple Choice

This section will contain 15 multiple choice questions on the contents of the whole course. Each question is worth 2 mark.

Example Questions

Ex Q1.	In a multilevel regression model, the random effects are:
A	Cluster level deviations from fixed intercepts and slopes
B	Outcome variables
C	The method used to find the maximum likelihood estimate
D	Cluster level intercepts and slopes

Ex Q2.	If we plot the group-level fitted lines from the following model, what would it look like? lmer(y ~ 1 + x + (1 | group))
A	Lots of lines that start in different places, and have different slopes
B	Lots of lines that start in the same place, but have different slopes
C	Lots of lines that start in different places, but for which the slope is the same
D	Just one line

Ex Q3.	A factor loading of 0.6 suggests that...
A	60% of the variance in the item is accounted for by the factor
B	0.60% of the variance in the item is accounted for by the factor
C	36% of the variance in the item is accounted for by the factor
D	0.36% of the variance in the item is accounted for by the factor

Ex Q4.	What property of a principal component analysis solution is lost when rotation is used?
A	Dividing eigenvalues by the number of items provides variance explained by the component
B	Sum of squared loadings on all components equals 1
C	Model identification
D	Sequential variance maximization

Section B: Written Answer & Interpretation

This section will contain 8 written answer & interpretation questions. Marks for each question are given at the start of the question. Some of the questions require you to interpret some R output for analyses we have been using in this course.

Example Question 1 [15 marks].

The 16 Personality Factor Questionnaire 5th Edition (16PF5) contains 185 items organized into 16 primary factor scales containing between 10 and 15 items each. The 16PF5 contains 15 primary personality scales, a 15 item Reasoning scale, and a 12 item Impression Management Scale.

A group of researchers were interested in studying whether the factor structure of the 16PF5 was stable when the response format was changed from the traditional 0-1-2 likert response scale, to a 1-10 sliding scale completed online. The results presented below show the analysis of two scales, Emotional Stability (C, 10 items) and Tension (Q4, 10 items).

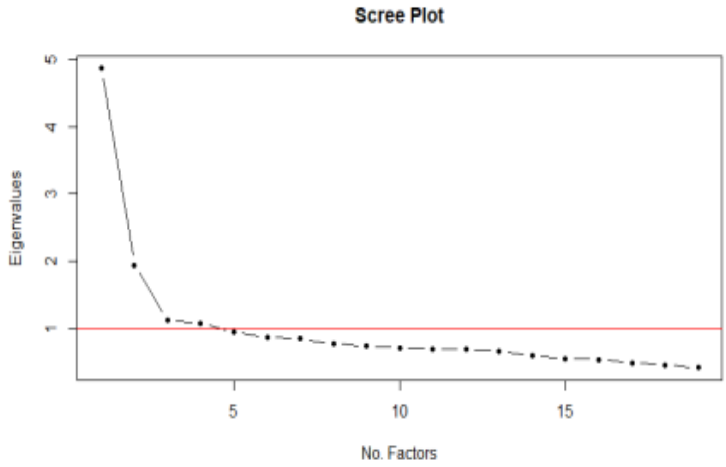
The scree plot for the analysis of the 20 items from these two scales is shown below.
The factor loading matrix from a minimum residual method exploratory factor analysis with oblimin (oblique) rotation is also shown.

Previous work based on large standardization samples has indicated that all items from both Emotional Stability (C, 10 items) and Tension (Q4, 10 items) load on to respective factors with loadings greater than 0.40, and that the factors have a low (< .20) correlation).
Based on the information provided, evaluate the plausibility of a two-factor solution, and the comparability of this factor solution to past results. Describe and provide a rationale for any next steps you would undertake in the analysis to explore the similarity of the solutions further.

```
Factor Analysis using method =  minres
Call: fa(r = efa2, nfactors = 2, rotate = "oblimin")
Standardized loadings (pattern matrix) based upon correlation matrix
       MR1   MR2   h2   u2 com
c1    0.44 -0.09 0.24 0.76 1.1
c2    0.52 -0.18 0.38 0.62 1.2
c3    0.38  0.03 0.14 0.86 1.0
c4    0.31  0.00 0.10 0.90 1.0
c5    0.52 -0.01 0.28 0.72 1.0
c6    0.66 -0.05 0.47 0.53 1.0
c7    0.68  0.12 0.41 0.59 1.1
c8    0.64 -0.02 0.43 0.57 1.0
c9    0.51 -0.04 0.27 0.73 1.0
c10   0.60  0.05 0.34 0.66 1.0
q4_1 -0.15  0.32 0.16 0.84 1.4
q4_2 -0.09  0.38 0.18 0.82 1.1
q4_3 -0.21  0.34 0.22 0.78 1.7
q4_4 -0.10  0.48 0.28 0.72 1.1
q4_5  0.02  0.65 0.41 0.59 1.0
q4_6  0.03  0.44 0.19 0.81 1.0
q4_7  0.12  0.63 0.35 0.65 1.1
q4_8 -0.07  0.53 0.32 0.68 1.0
q4_9 -0.22  0.42 0.30 0.70 1.5

                       MR1  MR2
SS loadings           3.20 2.25
Proportion Var        0.17 0.12
Cumulative Var        0.17 0.29
Proportion Explained  0.59 0.41
Cumulative Proportion 0.59 1.00

 with factor correlations of
      MR1   MR2
MR1  1.00 -0.42
MR2 -0.42  1.00
```



Example Question 2 [6 marks]

A researcher runs an exploratory factor analysis on two samples, extracting 4 factors in each case. In order to explore how similar the solutions are, they run a congruence analysis. The table below contains the results.

	MR1	**MR2**	**MR3**	**MR4**
MR1	0.78	0.22	0.30	0.10
MR2	0.30	0.95	0.17	0.28
MR3	0.10	0.24	0.33	0.43
MR4	0.20	0.11	0.97	0.35

Based on these results, what would you conclude about the replicability of the 4 factor solution, and briefly list what steps would you take next to find a replicable solution across samples?

Example Question 3 [10 marks]

A group of children are studying whether the 'spookiness' of their Halloween costumes is associated with how many treats are handed out during trick or treating, and whether this is different depending upon if the parents handing out candy have children of their own. The 25 children split up 400 houses in the neighbourhood, with each taking 16 houses.
For each house, they record the number of treats given and whether or not the household has children living there. The children's costumes were all rated on a spookiness scale by a team of independent raters.

Once the children have collected the data, they fit the model that is printed below. The spookiness rating of costumes is standardised.

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: treats ~ 1 + spookiness * hasKids + (1 + hasKids | child)
   Data: halloween

REML criterion at convergence: 2982.4

Scaled residuals:
     Min       1Q   Median       3Q      Max
-2.69183 -0.71069 -0.03452  0.68755  2.52688

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 child    (Intercept) 80.40    8.966
          hasKidsYes  38.37    6.194    -0.15
 Residual             83.97    9.163
Number of obs: 400, groups:  child, 25

Fixed effects:
                    Estimate Std. Error     df t value Pr(>|t|)
(Intercept)           34.080      2.147 26.825  15.871  3.7e-15 ***
spookiness             6.228      2.560 28.422   2.432   0.0215 *
hasKidsYes             3.955      1.754 23.004   2.255   0.0340 *
spookiness:hasKidsYes -3.608      2.100 24.693  -1.719   0.0982 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Provide an interpretation of each of:

- the fixed effects **[4 marks]**
- the random effects **[4 marks]**
- the relevance of the findings, considering the context of the study design and researchers' aims **[2 marks]**

Example Question 4 [6 marks]

Explain the differences between the 'complete pooling', 'no pooling', and 'partial pooling' approaches.