i Some guidelines:

- The exam consists of a combination of multiple choice questions and free-response questions. Note that the multiple choice questions have only one category which is the correct one.

- When answering questions that ask you to interpret the results, you can obtain partial credit even if you did not calculate the quantities correctly. As long as your interpretation is consistent with your calculations you will obtain at least partial credit.

- Read the questions carefully and write down if you do not understand something about a question. The responsible teacher will visit the exam site for a short period during the first two hours of the exam to answer clarifying questions.

Good luck!

/ Björn

1 Which of the following is a random variable?
Select one alternative:

○ The factor score.

○ The factor loading.

○ The error variance.

○ The observed sample mean.

Maximum marks: 2

2 The variance of a random variable X is 5. What is the variance of the random variable Y = 10X + 5?
Select one alternative:

○ 500

○ 5

○ 55

○ 50

Maximum marks: 2

3 Which of the following statements is always true for the classical true score model?
Select one alternative:

○ Higher difficulty means lower reliability.

○ The true score variance is always larger than the error score variance.

○ The true score is uncorrelated with the observed score.

○ Higher true score variance means higher reliability.

Maximum marks: 2

4 Assume that a single factor model is appropriate for a number of item scores and that coefficient alpha is equal to the reliability coefficient of the sum scores. Which of the following statements is a consequence of this?
Select one alternative:

○ The covariances between all item scores are equal.

○ The variances of the item scores are all equal.

○ The error variance is lower than the true score variance.

○ All items have the same difficulty.

Maximum marks: 2

5 Consider the factor model
$$X_j = \mu_j + \lambda_j F + \epsilon_j,$$
where F has variance 1. Consider an item with $\mu_j = 0.5$, $\lambda_j = 0.5$ and $\mathrm{Var}(X_j) = 0.85$. What is the variance of ϵ_j?
Select one alternative

○ 0.4

○ 0.5

○ 0.6

○ 0.1

Maximum marks: 2

6 Assume that a single factor model is appropriate for the analysis of several items. For a particular item, which of the following is evidence of item bias?
Select one alternative:

○ The reliability of the item score is different for men and women.

○ For men and women with remainder scores of 5, the means of the item score are different between men and women.

○ The means of the item score are different for men and women.

○ The variances of the item score are different for men and women.

Maximum marks: 2

7 For which of the following settings is equating necessary?
Select one alternative:

○ When using a cognitive scale to screen for dementia among people over the age of 60.

○ In a high-stakes college entrance exam meant to identify the top 25% most suitable students graduating from high school that year.

○ When using a scale to evaluate the level of depression among the teenage population.

○ In a low-stakes national reading exam, with purpose to assess the performance at the regional level over time in terms of fulfillment of learning outcomes.

Maximum marks: 2

8 Which of the following statements is most in line with the validity theory offered by the 2014 Standards for Educational and Psychological Testing?
Select one alternative:

○ The type of evidence required for validation depends on whether it is content, criterion, or construct validity that is to be established.

○ When a test score is used in multiple ways, each of these uses must be validated.

○ It is always necessary to consider all evidence categories in a validation study.

○ Construct validity is evaluated with factor analysis.

Maximum marks: 2

9 A test to measure reading literacy was developed and data were collected from a large group of students, along with the scores of an established test for interest in reading. You observed the following covariance matrix for the scores of the two tests, where X denotes the reading literacy test scores and Y denotes the interest in reading test scores:

$$\hat{\Sigma}_{X,Y} = \begin{pmatrix} \hat{\sigma}_X^2 & \hat{\sigma}_{X,Y} \\ \hat{\sigma}_{Y,X} & \hat{\sigma}_Y^2 \end{pmatrix} = \begin{pmatrix} 9 & 6 \\ 6 & 16 \end{pmatrix}.$$

Based on these observations, how would you characterize the relationship between reading literacy and interest in reading? State the assumptions made in the interpretations of the relationship.

Fill in your answer here

Maximum marks: 2

10 An item score X that could take integer values from 0 to 3 had the following probability distribution in a population of students:

P(X = 0) = 0.1
P(X = 1) = 0.1
P(X = 2) = 0.5
P(X = 3) = 0.3

a) What is the mean of X?

b) What is the mode of X?

c) What is the median of X?

d) What is the variance of X?

Maximum marks: 2

11 X and Y are two random variables with Var(X) = 2 and Var(Y) = 4.

a) Assume that X and Y are independent. Present the calculation of the variance of the random variable Z = X - Y.

Fill in your answer here

b) Assume instead that Cov(X, Y) = 0.5. Present the calculation of the variance of the random variable U = X + Y.

Fill in your answer here

Maximum marks: 2

12

Coefficient omega is defined as $\rho = \dfrac{\left(\sum_{j=1}^{J} \lambda_j\right)^2}{\sigma_Y^2}$.

For a three item test, the factor loadings were 1.2, 1.8 and 1.0, and the sum score variance was 20. Compute coefficient omega and interpret it. State the assumptions underlying the interpretation.

Fill in your answer here

7/15

Maximum marks: 2

13 The Standards for Educational and Psychological Testing state that validation practice can be guided by considering alternative hypotheses about interpretations of test-scores for particular uses. To aid generation of alternative hypotheses, the Standards state that it is useful to consider ways in which the test scores can be influenced by either (1) too much or (2) not enough.

A test is administered for the purpose of assessing the scholastic aptitude of students who are applying for university. The test includes assessment of logical reasoning, reading comprehension, vocabulary and the ability to understand charts and graphs. Provide **one example** of a way in which the test-scores might be influenced by too much, and **one example** of how the test-scores might be influenced by too little.

Fill in your answer here

| Format   ▾ | B | I | U | x₂ | x² | Iₓ | ⎘ | 📋 | ↰ | ↱ | ⟲ | ≔ | ☰ | Ω | ▦ | ✏ |

Σ | ⤢

Words: 0/200

Maximum marks: 2

14 For two tests X and Y, the linear equating function was estimated to be eq(Y) = 1.6 * X + 4. The cut score for passing test Y was determined to be 20. Give the cut score for pass in terms of the test X scores, based on the estimated equating function. Present how the result was obtained.

Fill in your answer here

| Format ▾ | B | I | U | X₂ | X² | Iₓ | ⧉ | 📋 | ↩ | ↪ | ⟲ | ⅑≡ | ⋮≡ | Ω | ⊞ | ✎ |
| Σ | ⤢ |

Words: 0/200

Maximum marks: 2

15 Item scores on a test of mathematics and a test of interest in mathematics were given to the same group of students. A two-factor model was estimated, yielding a model fit of GFI = 0.96, RMSEA = 0.04 and SRMR = 0.06. The correlation between the sum scores of the respective tests was 0.2 while the estimated factor correlation was 0.4. Give a reason why the factor correlation and the sum score correlation are not the same.

Fill in your answer here

Format | B *I* U x₂ x² | | | Ω | Σ

Words: 0/200

Maximum marks: 2

16 A bifactor model with one general factor and two subfactors (all factors independent) was estimated for an English test with two subdomains (reading and writing), yielding the following factor loading estimates:

Item	General	Reading	Writing
1	1	0.5	0
2	1	0.5	0
3	2	2	0
4	1	0	1
5	2	0	0.5
6	1	0	0.5

The model fit was judged to be acceptable.

In a previous study, the sum score was used. Based on the estimated factor loadings, would you recommend doing this?

Fill in your answer here

Words: 0/200

Maximum marks: 2

17 You have been asked to assist a group of English teachers to find the appropriate cut-score for a test of English reading proficiency. As part of the process, the test was piloted with a representative sample of the intended population and the results are available to you. In addition, the curriculum describes the expected level of English reading proficiency. Give a brief outline of how a standard-setting procedure could be used to find the cut-score for pass/fail on the English reading proficiency test.

Fill in your answer here

| Format ▾ | B | I | U | X₂ | X² | Iₓ | ⎘ | ⎙ | ↶ | ↷ | ⟳ | ⁝≡ | ⁞≡ | Ω | ⊞ | ✎ |
| Σ | ⤢ |

Words: 0/300

Maximum marks: 4

18 A scale is being developed to measure quality of life with the intended purpose to use the scale in national survey to identify which factors are associated with high quality of life in the population. The scale consists of Likert items. According to the underlying theory of quality of life, it is a multidimensional attribute. The theory also states that quality of life is expected to have differences based on gender.

Under these conditions, describe what evidence sources you want to consider in order to evaluate the validity of the scale scores for their intended purpose and describe the data you would like to analyze in doing so. Outline what results you would consider as evidence supporting the validity of using the scale scores in the national survey.

Words: 0/400

Maximum marks: 6

19 The following output was obtained from estimating a single factor model to five 4-category Likert scale items from an attitude scale.

Item	Factor Loading	Error variance
1	2.00	5.00
2	3.00	4.00
3	1.00	2.00
4	2.00	4.00
5	1.00	1.00

The model gave GFI = 0.92, SRMR = 0.08 and RMSEA = 0.07.

a) What validity evidence categories from the Standards for Educational and Psychological Testing are relevant in this analysis? (1p)

Fill in your answer here

b) Assume that a single factor model is appropriate. Which item contributes the most to the reliability of the sum score and which item contributes the least? Justify your answers. (2p)
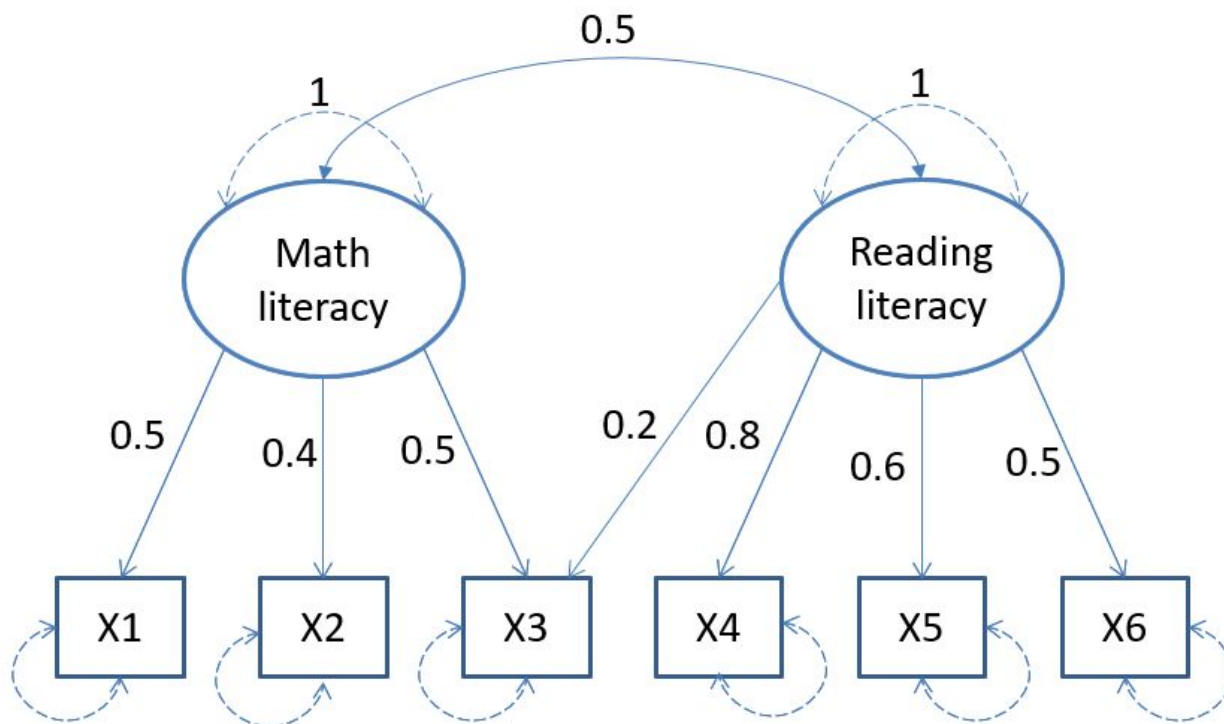
Fill in your answer here

c) From the description of the items above and the results of the estimated model, give two reservations against the use of the linear factor model in this case. (1p)

Fill in your answer here

Maximum marks: 4

20 Consider the graph below and answer the following questions.

a) What is the equation which describes the model for the item score X3? Write down the equation with an explanation of the parameters and variables included. (2p)

Enter text here

b) What is the covariance between item scores X3 and X4 according to the model? (2p)

Enter text here

Maximum marks: 4