Exam information

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 MC1H22

Which of the following is an estimate?

Select one alternative:

- The sample variance.
- The factor loading.
- The factor.
- The observed sample mean.

Maximum marks: 2

2 MC2H22

The variance of a random variable X is 2. What is the variance of the random variable $Y = 2X + 1$?

Select one alternative:

- 5
- 7
- 8
- 4

Maximum marks: 2

3 MC3H22

Which of the following statements is always true for an item score which conforms to a single factor model?

Select one alternative:

- The error score is uncorrelated with the factor score.
- The factor score variance is always larger than the error score variance.
- Higher difficulty means lower reliability.
- Higher error score variance means higher reliability.

Maximum marks: 2

4 MC4H22

Assume that a single factor model is appropriate for a number of item scores. Which of the following statements is a consequence of this?

Select one alternative:

- The covariances between all item scores must be equal.
- The reliability of the sum score can be expressed as a function of the factor loadings and the error variances.
- Coefficient alpha is equal to the reliability of the sum scores.
- The reliability of the sum score is above 0.8.

Maximum marks: 2

5 MC5H22

Consider the factor model $X_j = \mu_j + \lambda_j F + \epsilon_j$, where F has variance 1. For a specific item, $\mu_j = 1, \lambda_j = \frac{1}{2}$, and $\mathbf{Var}(X_j) = 1.$ What is the variance of ϵ_j?

Select one alternative:

- 3/4
- 1/4
- 0
- 1/2

Maximum marks: 2

6 MC6H22

Assume that a single factor model is appropriate for each of two groups (men or women) for the analysis of several items. For a particular item, which of the following is evidence of item bias?

Select one alternative:

- For men and women with the same factor score, the mean of the item score is higher for women than it is for men.
- The variance of the item score is higher for women than it is for men.
- The reliability of the item score is higher for women than it is for men.
- The mean of the item score is higher for women than it is for men.

Maximum marks: 2

7 MC7H22

For which of the following settings is equating necessary?

Select one alternative:

- When using a cognitive scale to screen for dementia in groups defined by their gender identity.
- In a low-stakes national reading exam, with purpose to assess the performance at the regional level over time.
- When using a scale to evaluate the level of anxiety among the teenage population.
- In a high-stakes college entrance exam meant to identify the top 25% most suitable students graduating from high school that year.

Maximum marks: 2

8 MC8H22

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:

- When carrying out a validation study, one must consider all evidence categories.
- When a test score is used in multiple ways, each of these uses must be validated.
- Evidence with regard to an external diagnostic criterion is always the gold standard validity evidence.
- The type of evidence required for validation depends on whether it is content, criterion, or construct validity that is to be established.

Maximum marks: 2

9 SR1H22

A scale to measure depression severity was developed and data were collected from a large group of students, along with the scores of an existing scale for satisfaction with life.

You observed the following covariance matrix for the scores of the two scales, where X denotes the depression severity scale scores and Y denotes the satisfaction with life scale scores:

$$\boldsymbol{\Sigma} = \begin{pmatrix} 10 & -7 \\ -7 & 10 \end{pmatrix}.$$

Based on these observations, how would you characterize the relationship between depression severity and satisfaction with life?

State the assumptions made in the interpretations of the relationship.

Fill in your answer here

Format

Words: 0

Maximum marks: 2

10 SR2H22

An item score X that could take integer values from 0 to 2 had the following probability distribution in a population of students:

P(X = 0)	P(X = 1)	P(X = 2)
0.2	0.6	0.2

Based on this information, answer the following questions. Only the answer is required.

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 SR3H22

X and Y are two random variables where $\mathbf{Var}(X) = 2$, $\mathbf{Var}(Y) = 3$ and $\mathbf{Cov}(X, Y) = 1$.

1. Calculate $\mathbf{Var}(Z)$, where $Z = X - Y$. Show your work.
2. Calculate $\mathbf{Var}(U)$, where $U = X + 2Y$. Show your work.

Fill in your answer here

Format | B I U x_2 x^2 I_x | Ω | Σ

Words: 0

Maximum marks: 2

12 SR4H22

Let m be the number of items on a test. For a five-item test, the common factor loading λ was 1 and the variance of the sum score Y was 10. Compute coefficient alpha

$\alpha = m\frac{\lambda^2}{\text{Var}(Y)}$

and interpret it. State the assumptions underlying the interpretation.

Fill in your answer here

Format B I U x₂ x² Ω Σ

Words: 0

Maximum marks: 2

13 SR5H22

The *Standards for Educational and Psychological Testing* (2014) state that it is useful to consider ways in which the test scores can be influenced by either (1) too much or (2) too little.

A three-domain test is administered for the purpose of measuring Norwegian 15-year-olds' ability to use their reading, mathematics and science knowledge and skills to meet real-life challenges. The test is a low-stakes test for the respondents since individual assessment is not of interest.

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

Words: 0

Maximum marks: 2

14 SR6H22

For two tests of reading comprehension, X and Y, the linear equating function was estimated to be $\text{eq}(Y) = 1.2X + 6$. The cut score for passing test Y was determined to be 30.
Give the cut score for pass in terms of the test X scores, based on the estimated equating function. Present and explain how the result was obtained.

Fill in your answer here

Format | B I U x_2 x^2 | I_x | | | Ω | |
Σ |

Words: 0

Maximum marks: 2

15 SR7H22

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 with correlated factors (one factor measured by the mathematics test items and the other by the interest in mathematics items) was estimated, yielding the model fit indices:

GFI	0.95
RMSEA	0.05
SRMR	0.06

The correlation between the sum scores of the respective tests was 0.2 while the estimated factor correlation was 0.5. Explain why there is a difference in the factor correlation and the sum score correlation in this context.

Fill in your answer here

Format

Words: 0

Maximum marks: 2

16 SR8H22

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

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

The model fit was judged to be satisfactory.

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

Fill in your answer here

Format

Words: 0

Maximum marks: 2

17 LR1H22

You have been asked to assist a group of teachers of Norwegian as a foreign language to find the appropriate cut-score for a test of Norwegian 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, an established framework describes the expected level of Norwegian 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 Norwegian reading proficiency test.

Fill in your answer here

Format

Words: 0

Maximum marks: 4

18 LR2H22

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

With this information in mind, do the following:

- Describe what evidence sources you want to consider in order to evaluate the validity of the scale scores for their intended purpose
- Describe the data you would like to collect to conduct the validity study
- Describe the analyses you would do in the validity study
- Outline what results you would consider as evidence supporting the validity of using the scale scores in the national survey

Fill in your answer here

Format

Words: 0

Maximum marks: 6

19 LR3H22

The following output was obtained from estimating a single factor model to five 4-category Likert scale items from a scale measuring the environmental awareness of 15-year olds in Norway.

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

The residual correlation matrix was

$$\boldsymbol{\Sigma}_{\text{res}} = \begin{pmatrix} \mathbf{0.000} & & & & \\ \mathbf{0.026} & \mathbf{0.000} & & & \\ \mathbf{0.017} & \mathbf{-0.035} & \mathbf{0.000} & & \\ \mathbf{-0.014} & \mathbf{0.072} & \mathbf{-0.019} & \mathbf{0.000} & \\ \mathbf{-0.025} & \mathbf{-0.039} & \mathbf{0.020} & \mathbf{0.009} & \mathbf{0.000} \end{pmatrix}.$$

Address the following in your response:

1. What validity evidence categories from the Standards for Educational and Psychological Testing are relevant in this analysis? (1p)
2. Based on your appraisal, does the single factor model fit well?
3. Assume that a single factor model is appropriate for the analysis of the five item scores. Which item contributes the most to the reliability of the sum score and which item contributes the least? Justify your answers. (1p)
4. From the description of the items above and the results of the estimated model, give **one reservation** against the use of the linear factor model in this case. (1p)

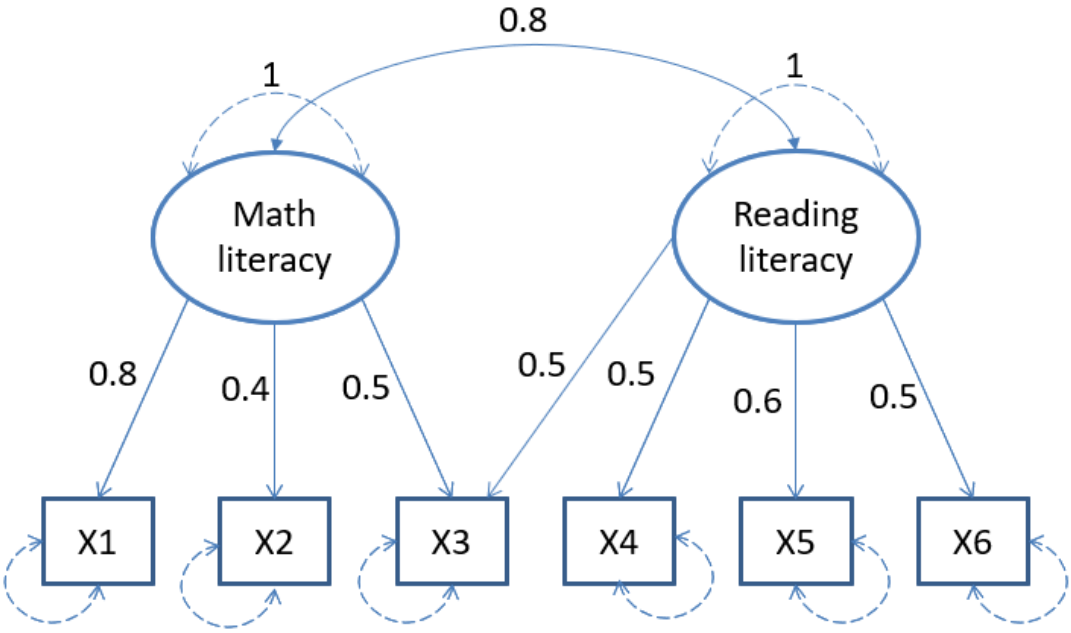
Fill in your answer here

Words: 0

Maximum marks: 4

20 LR4H22

A multiple factor model is illustrated in the graph below. The latent variables and the observed variables are all standardized.



Answer the following questions based on the graph.

1. 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)
2. What is the covariance between item scores **X3** and **X4** according to the model? Show your work and explain the steps taken. (2p)

Fill in your answer here

Words: 0

Maximum marks: 4