

**UiO : Universitetet i Oslo**

CANDIDATE

**184118**

TEST

# MAE4011 1 Principles of Measurement

Subject code	MAE4011
Evaluation type	Individuell skriftlig prøve
Test opening time	20.12.2022 09:00
End time	20.12.2022 13:00
Grade deadline	--
PDF created	26.12.2022 19:33
Created by	Tony Clifford Austin Tan

**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:

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

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

Words: 0

Unanswered.

**11 SR3H22**

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

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

Fill in your answer here

Words: 0

Unanswered.

**12 SR4H22**

Let  $m$  be the number of items on a test. For a five-item test, the common factor loading  $\lambda$  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**

Words: 0

Unanswered.

**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**

Words: 0

Unanswered.

**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**

Words: 0

Unanswered.

**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**

Words: 0

Unanswered.

**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**

Words: 0

Unanswered.

**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**

You could have a number of different tests within the reading proficiency test, which could give different scores based on their abilities. First you would have to make a test that captures their overall reading proficiency. You could test the person's ability to recall details from a text, give an overall summary of a text, a grammar part where the person could spot the miss-spelled words in a text etc. Based on the different parts of the tests there can be given a score to the student, which can be scaled so that each part weighs the same. Then they could set the minimum requirement to be able to pass (for example a 40% overall score). To ensure the person has an overall good enough ability within reading proficiency, they could say that in addition to being able to get a 40% overall score the scores need to be equally divided based on the different parts of the test.

Words: 157

---

Answered.

**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**

To be able to evaluate the validity of the scale scores for their intended purpose i would consult with a panel of experts who could give some pointers on what to include in the scale. I would spend time gathering relevant information and a proper list of sources that in depth explores the subject of life satisfaction. I would look in to already existing measures for life satisfaction and based on all the information gathered i would make a first draft for the measurement. I would then gather a cognitive lab to ensure that the questions are understood the way their intended and that the scale is measuring whats intended. If there is a already existing measurement that is already acknowledged within the research area, i would gather a test group to fill out forms for both the already existing test and the newly developed to check the correlation between the test scores. In the data i would make sure to collect data that measures satisfaction within different aspects of life; school/work, social life, physical health, economic situation, mental health etc. I would make sure to ask some questions that ask for overall satisfaction in addition to questions within specific categories to ensure that we get a proper overall view. Within each of the categories i would add a question asking how important this part of the persons life is to them, to ensure that for example if a person understands the question about economic situation as a question about their income, and has a below-average income, and therefor gives a low mark on the topic it can be "canceled" out if the person are satisfied with their income. In the validity study i would check if a single factor model fits, and check the factor loadings in addition to the error and the variance to see the items relationship to the construct. I would the calculate the omega coefficient to see the internal validity. If the single factor model does not fit, i would make a multiple factor model and check the coefficient alpha (all though if the single factor model does not fit, i would probably look into why since its supposed to be a single construct measurement which should fit in a single factor model).

Words: 377

Answered.

## 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

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

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)

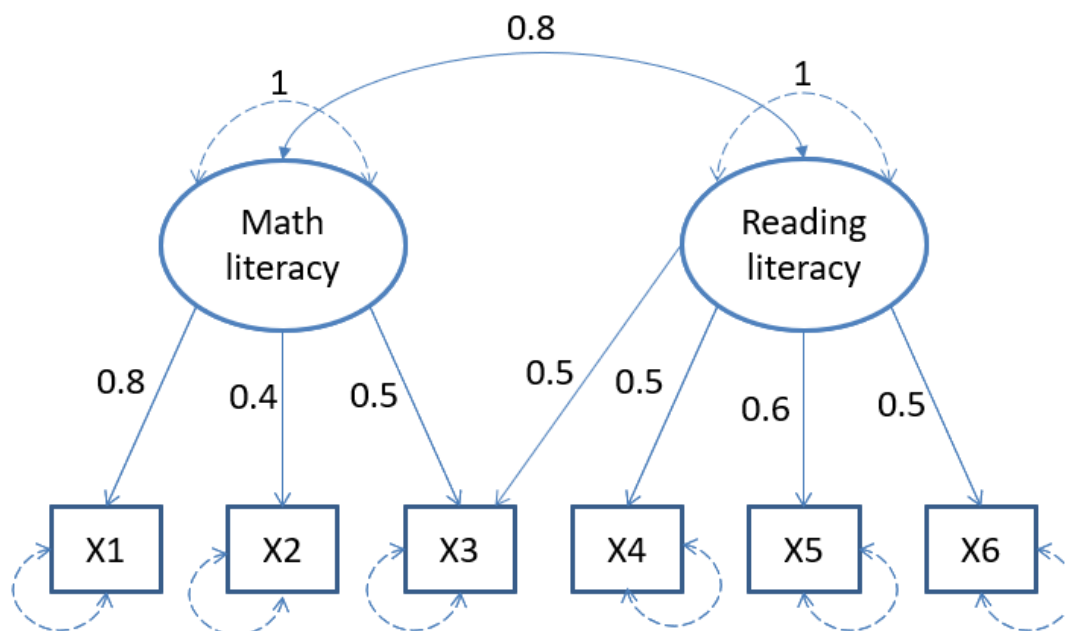
Fill in your answer here

Words: 0

Unanswered.

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

Unanswered.