Principles of Measurement - Assignment 2

November 11, 2021

Instructions

- The assignment should be completed individually.

- Write your answers in a word processor.

- Upload your assignment through the submission system in canvas.

- The deadline for handing in the assignment is November 26 18.00.

Exercise 1

Consider the scale concerning environmental awareness, which has seven items. Data were collected at two time points and ten complete observations were obtained. The data can be found in the file `MAE4011H21A2.RData`. Define Likert-scale item scores and consider the following analyses.

a (max 100 words)

Estimate the test-retest reliability of the summed item scores. Comment on the results.

b (max 100 words)

Estimate coefficient alpha at each of the two time points. Compare the obtained estimates to the estimated test-retest reliability. Comment on the results.

Exercise 2

You should discuss one of the following journal articles:

- Ching, E. W. K., Chiang, T. P., Wong, J. O., Siu, B. W., Lui, S. S., & Cheung, E. F. (2021). Validation of the Chinese Dimensional Anhedonia Rating Scale in Depressed Patients in Hong Kong. *Psychological Test Adaptation and Development.* `http://doi.org/10.1027/2698-1866/a000012`

- Freitas, S., Simoes, M. R., Alves, L., & Santana, I. (2013). Montreal cognitive assessment: validation study for mild cognitive impairment and Alzheimer disease. *Alzheimer Disease & Associated Disorders*, 27, 37-43. `https://doi.org/10.1097/WAD.0b013e3182420bfe`

- Hartmaier, S. L., Sloane, P. D., Guess, H. A., Koch, G. G., Mitchell, C. M., & Phillips, C. D. (1995). Validation of the minimum data set cognitive performance scale: agreement with the mini-mental state examination. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences, 50*, M128-M133. `https://doi.org/10.1093/gerona/50A.2.M128`

- Wann, D. L. (1995). Preliminary validation of the sport fan motivation scale. *Journal of Sport and Social issues, 19(4)*, 377-396. `https://doi.org/10.1177/019372395019004004`

- Wickramasinghe, V. P., Cleghorn, G. J., Edmiston, K. A., Murphy, A. J., Abbott, R. A., & Davies, P. S. (2005). Validity of BMI as a measure of obesity in Australian white Caucasian and Australian Sri Lankan children. *Annals of human biology, 32(1)*, 60–71. `https://doi.org/10.1080/03014460400027805`

Based on your input, we have allocated the papers as follows:

Name	Article
Lucy	Ching et al. (2021)
Xiao	Freitas et al. (2013)
Yuriko	Wann (1995)
Wonseok	Ching et al. (2021)
Sverre	Wann (1995)
Tore	Wickramsinghe et al. (2005)
Amy	Hartmeier et al. (1995)
Markus	Wickramsinghe et al. (2005)
Erika	Wann (1995)
Munyaka	Wann (1995)
Raphael	Ching et al. (2021)
Jianan	Freitas et al. (2013)

The task (max 350 words)

Briefly outline what kind of validity evidence or validation procedure is presented in your designated paper. Discuss the validity evidence resulting from the approach used or presented in relation to the test score usage in the particular context. Specifically, your answer should:

  a Identify the claims made in the article that are of relevance with respect to the test score interpretations.

  b Categorize the validation procedures outlined in the article with respect to the evidence categories of the Standards for Educational and Psychological Testing.

  c Evaluate the extent to which the originally stated interpretations of the test scores are valid, considering the evidence that was collected as part of the study.

Exercise 3

You have been assigned **one** of the scenarios below as indicated in the table here. Consider your specific scenario and answer the task at the end.

Scenario	Name	Data set
1	Lucy	1
2	Amy	2
3	Erika	3
4	Tore	4
1	Raphael	5
2	Jianan	6
3	Yuriko	7
4	Sverre	8
1	Xiao	9
2	Markus	10
3	Munyaka	11
4	Wonseok	12

Scenario 1

A research group has developed a five-item scale with 7-category integer scored items (variables SWLS1 to SWLS5) that they aim to use in a Norwegian national survey as a measure of satisfaction with life. Validity evidence with respect to content has been established by considering expert opinion. According to the prevailing theoretical model, satisfaction with life is a unidimensional construct that is moderately positively associated with job satisfaction and strongly negatively associated with narcissistic personality disorder.

The researchers have collected item scores from 152 undergraduate students at the Department of Psychology at the University of Oslo. In addition, data from two established scales measuring job satisfaction (variables JS1 to JS5; reliability 0.83) and narcissistic tendency (variables NS1 to NS5; reliability 0.80) were collected.

Scenario 2

A research group has developed an algebra test with five separate parts scored as integers between 0 and 6 (variables Alg1 to Alg5). The test is aimed to be used as a diagnostic tool among Norwegian grade 7 students, to identify students in need of additional help. Validity evidence with respect to content has been established by considering expert opinion. According to the prevailing theoretical model, algebra performance is a unidimensional construct

that is strongly positively associated with reading proficiency and weakly positively associated with interest in mathematics.

The researchers have collected algebra item scores from 210 grade 6 to 8 students in the Oslo region. In addition, item scores from scales measuring reading proficiency (variables R1 to R5; reliability 0.82) and interest in mathematics (variables M1 to M5; reliability 0.78) were obtained from the same students.

Scenario 3

A test intended to measure self-determination was developed for use on research staff in the academic sector. The test consists of five Likert-scale items (variables SD1 to SD5), where respondents indicate the degree to which they agree or disagree with statements believed (based on the construct-theory) to reflect the target attribute. A panel of experts judged the test-content to be in alignment with the construct-theory. According to theory, the subjective experience of autonomy construct is unidimensional, and is expected to exhibit a strong relationship (correlation in the range of .7 - .8) with the job-satisfaction construct and a weak relationship with the extraversion personality trait.

To evaluate the validity of interpretations of the self-determination scale scores, the test was administered to a sample of 93 permanent and temporary staff at the University of Oslo. Shortened versions of previously validated versions of tests measuring the job-satisfaction (variables JS1 to JS5; reliability 0.82) and extraversion (variables EV1 to EV5; reliability 0.78) constructs were also included as part of the data-collection.

Scenario 4

A group of researchers translated a scale for problem-solving with 5 items, scored as integers from 0 to 4 (variables PS1 to PS5). In the original language, the scale for problem-solving was developed together with a scale measuring creativity, and a moderately strong association existed between the two constructs. Additionally, it has been shown that problem-solving performance correlates strongly with spatial reasoning skills from previous research.

To validate the translated scale for problem-solving, the researchers recruited 106 students to participate in a validation study. The researchers

included a previously validated translation of the scale for creativity (C1 to C5; reliability 0.88), and a measure of spatial reasoning skills (SR1 to SR5; reliability 0.82)

The task (max 700 words)

Investigate the evidence for and against the interpretation of the sum scores of the life satisfaction, algebra, self-determination or problem-solving tests as reliable measures of the constructs in the population where the scores are meant to be used. Indicate what additional aspects should be considered in future studies. The data are in the lists `Scenario1`, `Scenario2`, `Scenario3`, and `Scenario4` in the file `MAE4011H21A2.RData` and you should use the list and list entry that corresponds to the Scenario and Data set numbers given in the table above.