JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

Students' effort and motivation in PISA 2015 – 2022 in Finland

Heli Kauppinen
Project Researcher

Arto Ahonen
Senior Researcher, National Project Manager of Finland for PISA

Finnish Institute for Educational Research

Background

- PISA is a low-stakes test for students, meaning there are no consequences for poor performance or non-participation.
- PISA does not require any preparation and is designed to assess what students can do with the knowledge and skills they have learned in a typical school day.
  -> Different from a regular school test, in which students prepare and present their best knowledge in specific fields.
- The students will not receive feedback on their performance during or after the test. These factors can undermine student's motivation and willingness to do their best in the PISA test.
- In general, students' efforts on the PISA test in Finland have declined from 2015 to 2018 and again in 2022.

Effort and test motivation

- Expectancy-value theory assumes that effort is the outcome of expectancy and is related to test performance (Penk & Richter 2017).
- The effort put into a test has been found to improve the outcomes, even though the test does not have an impact on the students themselves. (Zilbelberg, Finney, Marsh & Anderson 2014; Cole, Bergin & Whittaker 2008).

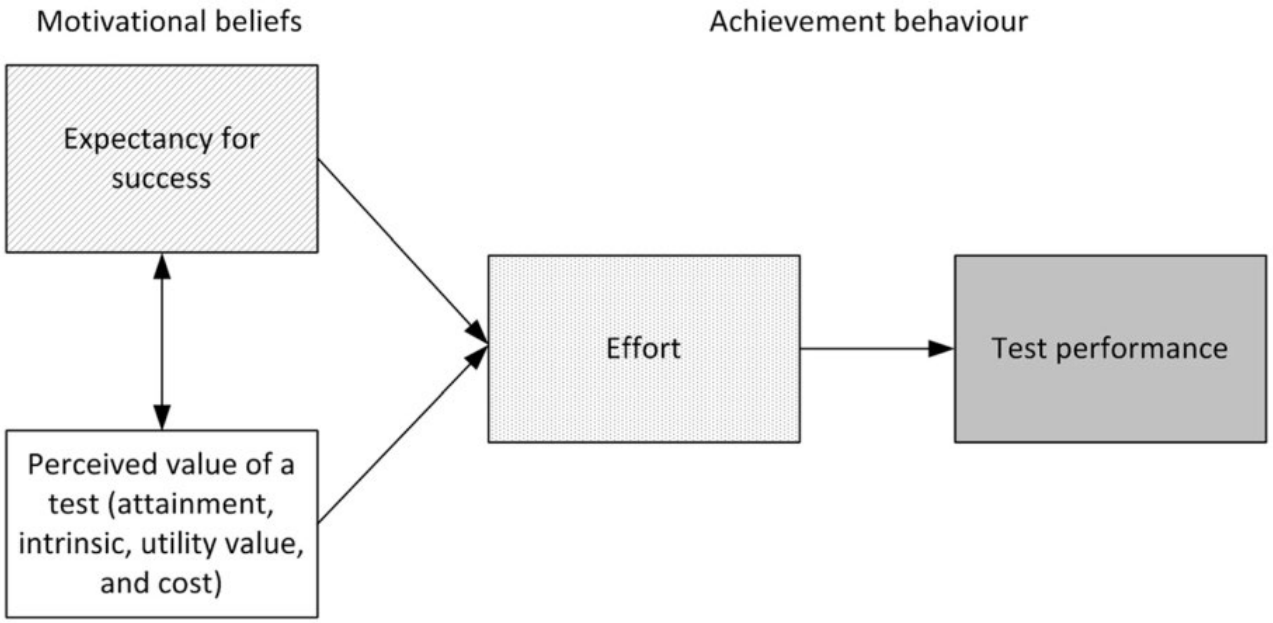
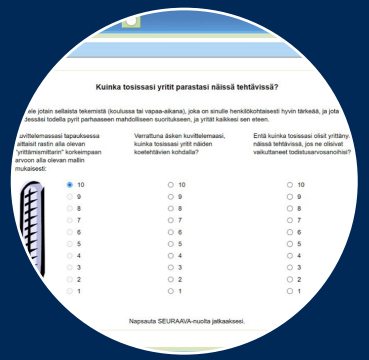


Fig. 1 Expectancy-value theory in the context of test-taking motivation (adapted from Eccles and Wigfield 2002; Wigfield and Eccles 2000)

How student effort and motivation is measured in PISA?



Self-reported measures

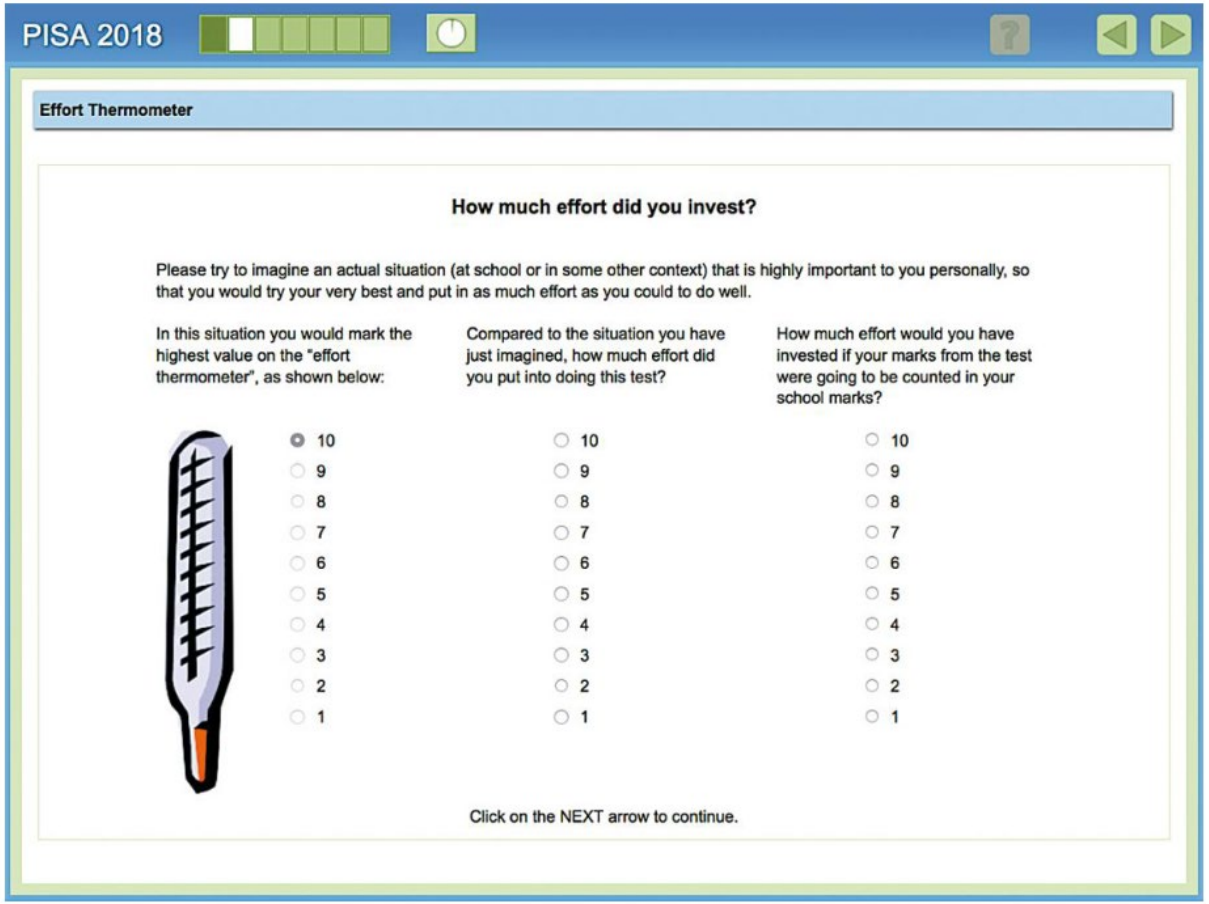
- Effort Thermometer (at the end of the cognitive part of test)
- National item scale (at the end of the questionnaire part)

PISA Log data (CBA)

- Response-time
- Time of first action
- Number of visits, short visits and actions

Effort Thermometer

- ”How much effort did you put into doing this test?
- ”How much effort would you have invested if your marks from the test were going to be counted in your school marks?”
- In 2022 study, in Finland the average for the first question was 7.7 and second 9.1.
- In 2018 the averages were 8.0 and 9.3.
- In the subset of PISA data used in our analysis, the average for the first question was 7.8 in 2022 and 8.0 in 2018.



Measuring effort with national item scale

- Students' lowered motivation towards school has been a topic in education policy discussions both in Finland and internationally.
- We have used the Validated Student Opinion Scale (SOS), developed by Wolf and Smith (1995) and Sundre (1999), as an optional national question in STQ.
- Our aim is to investigate **how students' effort and motivation in PISA have changed** between 2015 and 2022 and **how this change is related to changes in mathematics proficiency** in Finland.

- **The SOS effort scale (Likert 1-4)**
  - I was motivated to do my best in the PISA test.
  - I made a good effort in the PISA test.
  - It was important for me to do well in the PISA test.
  - I worked with the problems without giving up even though some of them were difficult.
  - It means a lot to me to do well in the PISA test.
  - I did my best in the PISA test.

Data and measures

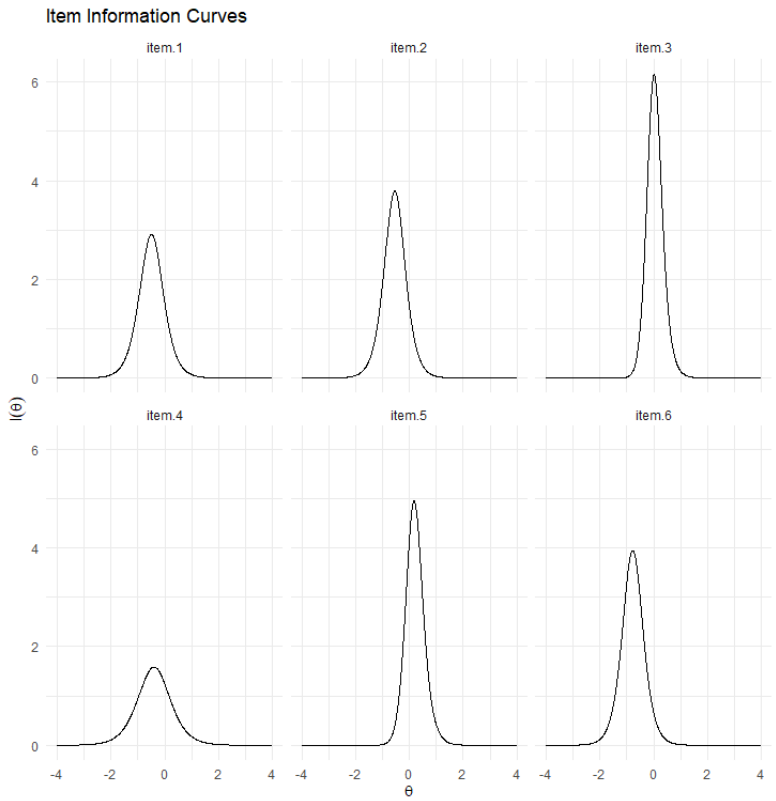
- Finnish PISA datasets 2015 (N = 5882), 2018 (N = 8243) and 2022 (N = 10 239).
- **Students with no responses to the national items were excluded from the dataset. After omitting missing data, the total N = 16 977.**
- To measure students' effort and test motivation in PISA, the national effort scale items were recoded (1 = agree, 0 = disagree).
- A new effort scale was calculated implementing IRT (item response theory) approach, in more detail, 3PL-model. (see Next slide)
- The relationship between effort and mathematics achievement was modeled using survey-weighted generalized linear models with the R package *survey* (Lumley 2004). Achievement-related models were run across 10 mathematics plausible values, after which the estimates and standard errors were obtained using Rubin's Rules (RR, see Rubin 1987).

Table 1. The proportions of girls and boys in the data subset used in the analysis

	Total N	Girls N (std. Weighted %)	Boys N (std. Weighted %)
PISA 2015	5152	2531 (49%)	2621 (51%)
PISA 2018	4817	2410 (50%)	2407 (50%)
PISA 2022	7008	3447 (49%)	3561 (51%)
Total	16977	8388 (49%)	8589 (51%)

Brief Summary of 3PL Model Analysis

- Data was recoded to binary (1 = agree, 0 = disagree)
- Datasets were combined over 2015, 2018 and 2022 prior to modeling.
- Model was fitted using R with *mirt* package (Chalmers 2012).
- **3PL Model**: Three-parameter logistic model. Parameters:
  - **Discrimination (a)**: How well an item differentiates between individuals with different levels of the latent trait (effort).
  - **Difficulty (b)**: The level of the effort at which the item has a 50% chance of being agreed.
  - **Guessing (g)**: The probability of "guessing the right answer". In other words, probability of student agreeing on item without truly reflecting the latent trait (effort).
- Proportion of total variance: 83%
- The effort scores for the students were calculated using WLE estimates (Warm, 1989).



	a	b	g	u
I was motivated to do my best in the PISA test.	3.414	-0.501	0	1
I made a good effort in the PISA test.	3.892	-0.543	0	1
It was important for me to do well in the PISA test.	5.17	-0.003	0.042	1
I worked with the problems without giving up even though some of them were difficult.	2.518	-0.408	0	1
It means a lot to me to do well in the PISA test.	4.592	0.165	0.032	1
I did my best in the PISA test.	3.973	-0.784	0	1

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

Research questions

1. How has the effort and test motivation of Finnish students changed in PISA test during 2015, 2018 and 2022?
2. What is the relationship between effort, test motivation and mathematics achievement in PISA? How has this relationship changed from 2015 and 2018 to 2022?

1. How have Finnish students' effort and test motivation changed in PISA test during 2015, 2018 and 2022?

Figure 1. Distribution of responses by national effort items in PISA 2015 – 2022

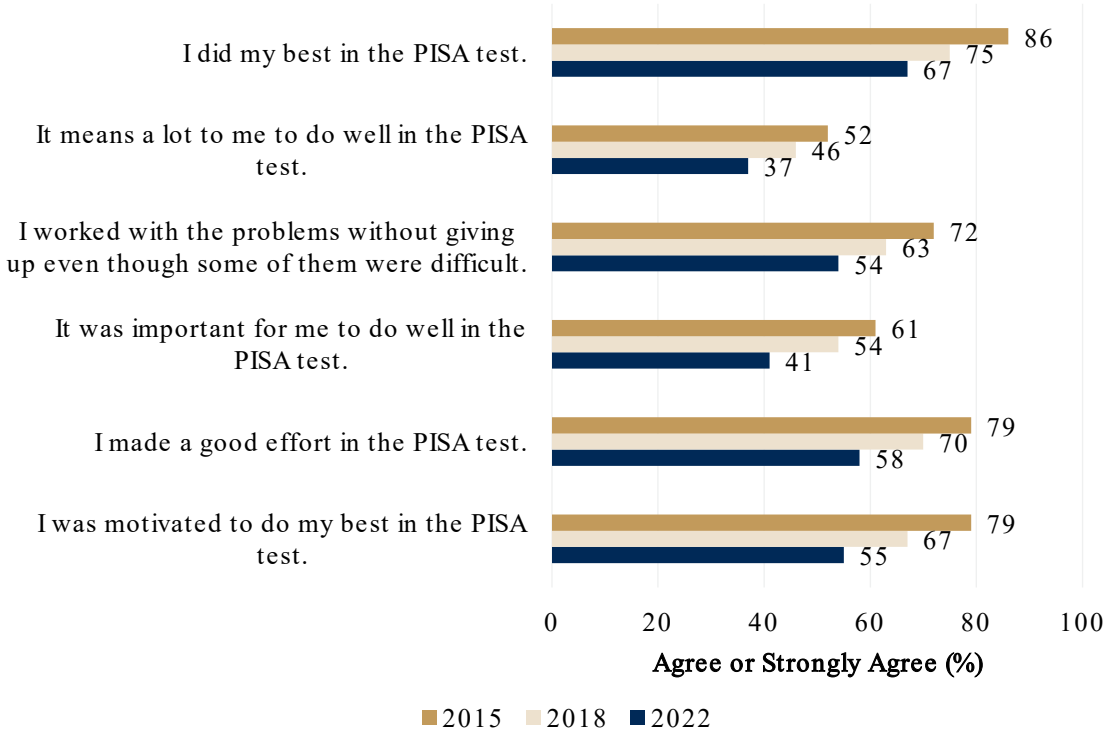
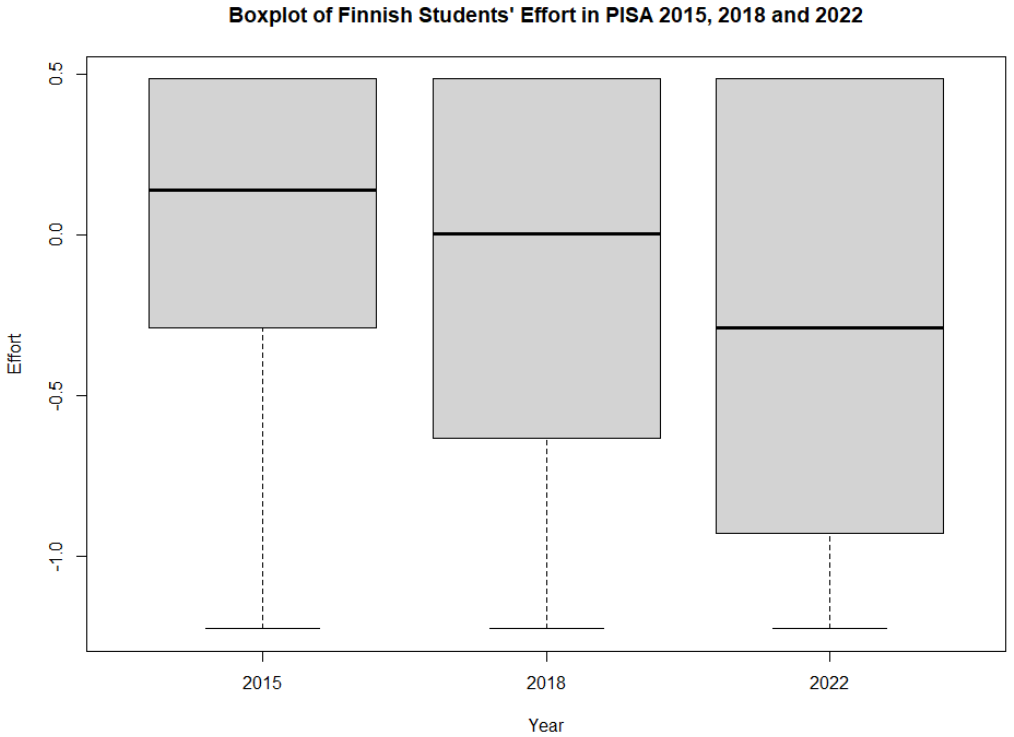
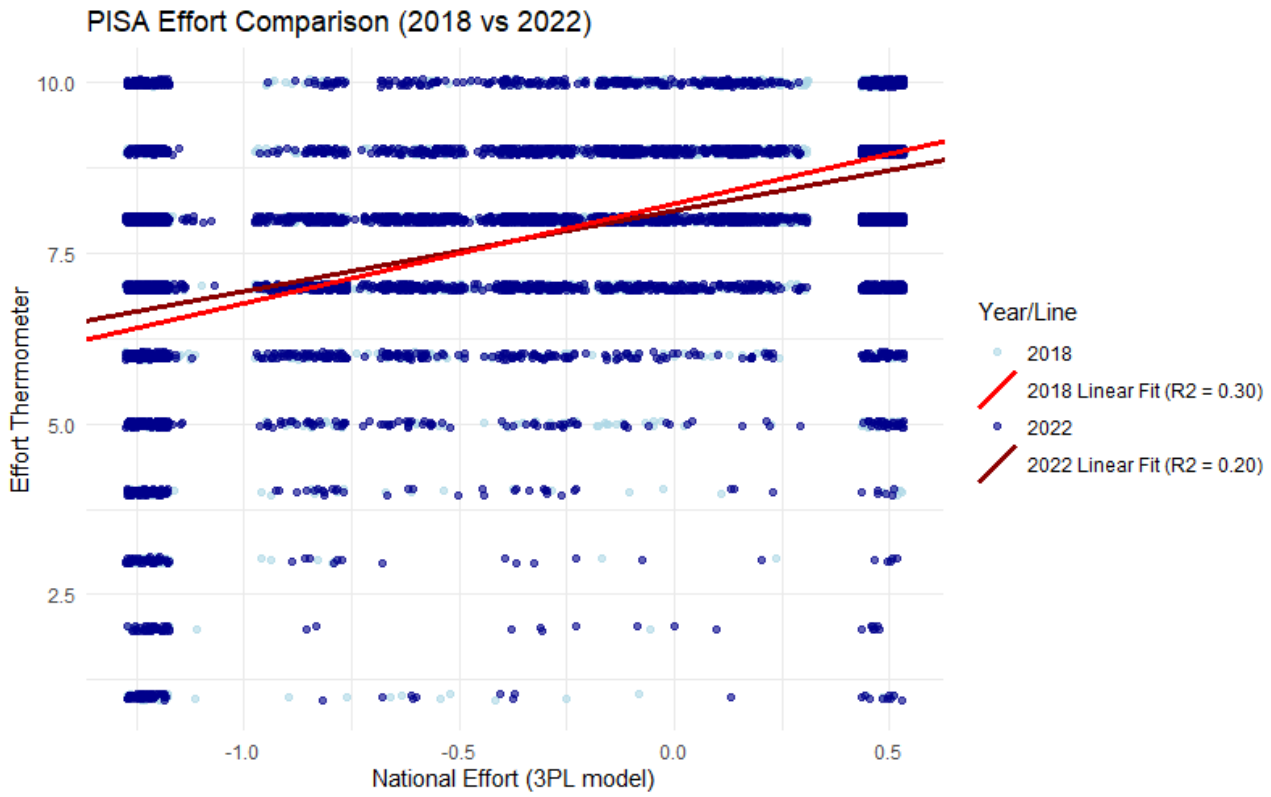


Figure 2. Boxplot of Finnish Students' Effort in PISA 2015, 2018 and 2022



How are the results aligned with the Effort Thermometer (PISA 2018 & PISA 2022)?

Figure 3. Scatterplot of National Effort Scale and Effort Thermometer in Finnish PISA 2018 and 2022



In 2018, the correlation between Effort Thermometer and National Effort Scale was 0.54. In 2022, the correlation was 0.44.

Table 2. Summary of results when mathematics achievement is explained by the National Effort Scale (PISA 2015, 2018 & 2022)

	Estimate
Intercept	516.3**
PISA 2018	0.4
PISA 2022	-17.0**
National Effort	
PISA 2015	36.0**
PISA 2018	33.3
PISA 2022	28.2*

Table 3. Summary of results when mathematics achievement is explained by international Effort Thermometer (PISA 2018 & 2022)

	Estimate
Intercept	513.7
PISA 2022	-18.9*
Effort Thermometer	
PISA 2018	23.9**
PISA 2022	23.6

2. What kind of relationship effort and test motivation have to the mathematics achievement in PISA? How has this relationship changed in 2015, 2018 and 2022?

- The regression coefficient of student effort has decreased year after year.
  - In PISA 2015, when comparing two students, the student with one-unit higher national effort scored on average 36 points higher.
  - In PISA 2022, the according difference was 28 points on average.

Are the results of the National Effort Scale in line with the Effort Thermometer?

- The overall results are similar for both models, but the coefficient of determination is a bit bigger in Effort Thermometer model (R2 = 0.09) than in the National Effort Scale model (R2 = 0.07).

Keypoints

Effort and motivation in PISA test seem to be declining in Finland

- As measured by the national effort items analyzed using the IRT model approach, effort and test motivation in PISA have decreased from 2015 to 2022.
- The national average of the effort thermometer has also declined.

The relationship between effort and performance in mathematics has become weaker

- Overall, the coefficient of determination was smaller when using national effort scale.
- The linear relationship between effort, test motivation and mathematics achievement in PISA has become weaker.
- Still, the effort in PISA test has a positive relationship with mathematics achievement.

What next?

- The national effort scale brings a new perspective of students' test motivation and effort in PISA test.
  - Effort thermometer seem not to measure the exactly the same thing.
- It could be beneficial to utilize the log data and combine it with the self-reported effort scales to enhance their accuracy.

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

https://www.jyu.fi/en
https://ktl.jyu.fi/en

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

References

Chalmers, R. P. (2012). *mirt: A multidimensional item response theory package for the R environment*. Journal of Statistical Software, 48, 1-29.

Cole, J. S., Bergin, D. A., & Whittaker, T. A. (2008). Predicting student achievement for low stakes tests with effort and task value. Contemporary Educational Psychology, 33(4), 609-624.

Lumley, T. (2004). Analysis of Complex Survey Samples. *Journal of Statistical Software*, *9*(8), 1–19. https://doi.org/10.18637/jss.v009.i08

Penk, C., & Richter, D. (2017). Change in test-taking motivation and its relationship to test performance in low-stakes assessments. *Educational Assessment, Evaluation and Accountability, 29*(1), 55–79. https://doi.org/10.1007/s11092-016-9248-7

Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons.

Sundre, D. L. (1999). Does examinee motivation moderate the relationship between test consequences and test performance? Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, *54*(3), 427-450.

Wolf, L. F., & Smith, J. K. (1995). The Consequence of Consequence: Motivation, Anxiety, and Test Performance. Applied Measurement in Education, 8(3), 227–242. https://doi.org/10.1207/s15324818ame0803_3

Zilberberg, A., Finney, S. J., Marsh, K. R., & Anderson, R. D. (2014). The role of students' attitudes and test-taking motivation on the validity of college institutional accountability tests: a path analytic model. International Journal of Testing, 14(4), 360–384. https://doi.org/10.1080/15305058.2014.928301