

**Measuring Change And
Different Response Format
Effects In Large Scale
Educational Testing With
LLRA**

Rusch, Thomas; Dobrovits, Ingrid; Gatterer, Birgit; Hatzinger, Reinhold

Published: 01/01/2011

Document Version
Unknown

[Link to publication](#)

Citation for pulished version (APA):

Rusch, T., Dobrovits, I., Gatterer, B., & Hatzinger, R. (2011). *Measuring Change And Different Response Format Effects In Large Scale Educational Testing With LLRA.*

Measuring Change And Different Response Format Effects In Large Scale Educational Testing With LLRA

- 1 Background And Goal
- 2 Study Design
- 3 Linear Logistic Models With Relaxed Assumptions
- 4 Results
- 5 Conclusion And Outlook

This is joint work with Ingrid Dobrovits, Birgit Gatterer and Reinhold Hatzinger.

- WU is the biggest European business school
- Each year about 4500 new students enroll at WU
- There is an entry phase that has to be passed to continue with further courses (STEP)
- The STEP is followed by courses that form a Common Body Of Knowledge (CBK)
- Most exams in these two parts employ multiple choice response formats
- A small number of courses still uses an open format which is impractical

- A major exam in the CBK module is Accounting, Management and Control
- Examination of students happened traditionally with open answer formats
- Last year multiple choice questions have been developed which shall eventually substitute the traditional format
- There is some concern that the MC format may be less suited to assess students' abilities however
- The Institute of Business Education started a study to investigate the possible effects of changing the response format

Items and Scales

Overall 12 items were used to assess students' skills in the subject.
 The items should form three content-related scales.

- Scale 1 (low value fixed assets): Items 1, 2, 3, 4
- Scale 2 (dubious claims): Items 5, 7, 9, 11
- Scale 3 (unmeetable claims): Items 6, 8, 10, 12

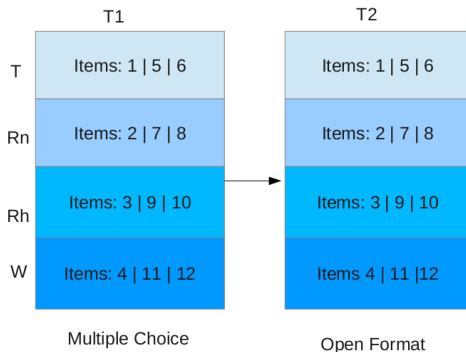
For each scale, the items stand for a specific necessary aspect of skill in that topic

- Transfer (T)
- Low routine task (Rn)
- High routine task (Rh)
- Knowledge (W)

The study had the following goals

- Establish if there is equivalency between the multiple choice and the open format
- Identify possible differences in the answer formats
- Explore the possible impact of response format change on different groups
- Assess “fairness” of the questions
- Check the scale assumptions

The (simplified) study design was

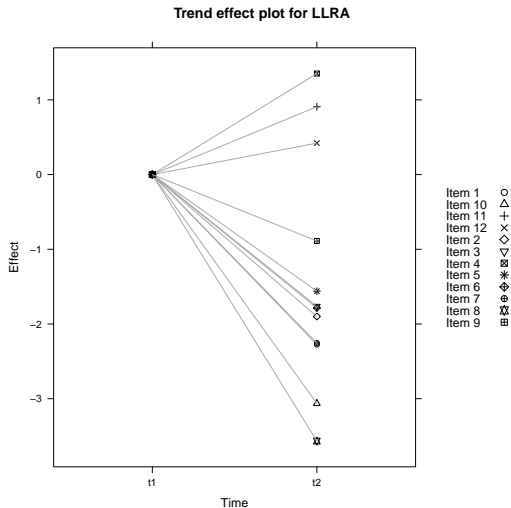


The design was not well suited to conclusively answer all questions but by treating each item as its own dimension we might still get some answers.

To this end we used Linear Logistic Models With Relaxed Assumptions (LLRA) (Fischer, 1989; Fischer, 1995).

- They allow to analyse categorical repeated measurement data
- They require neither unidimensionality of items nor distributional assumptions about the latent trait
- They allow to contrast treatment and/or subject covariate effects
- They provide a framework for testing assumptions on (item) dimensionality, trend and covariate effects
- Change is modelled as a linear function of trend effects and covariate/treatment main and interaction effects
- Due to conditional maximum likelihood estimation (CML), inference on effect parameters is completely independent of the trait parameters

Results: Change Over Time - I



Results: Change Over Time- II

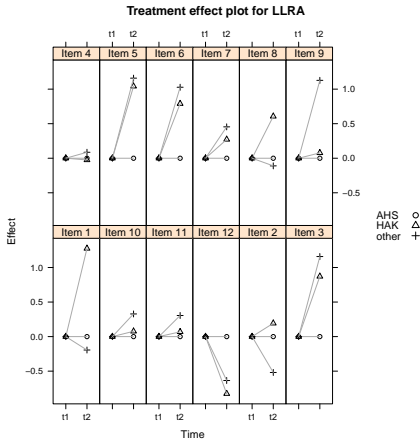
Item	Estimate	Std.Error	lower.CI	upper.CI
I1	-2.276	0.219	-2.705	-1.847
I2	-1.900	0.246	-2.382	-1.418
I3	-1.762	0.226	-2.205	-1.320
I4	1.352	0.205	0.951	1.754
I5	-1.561	0.189	-1.931	-1.192
I6	-1.785	0.212	-2.201	-1.370
I7	-2.255	0.271	-2.787	-1.723
I8	-3.567	0.383	-4.319	-2.816
I9	-0.889	0.179	-1.240	-0.538
I10	-3.064	0.295	-3.643	-2.485
I11	0.908	0.175	0.565	1.250
I12	0.421	0.140	0.146	0.696

We looked at effects of the following variables

- **Gender:** No group effect, trend effect does not change
- **Native speaker:** No group effect, trend effect does not change
- **Positive grade on MC:** No group effect, Item 12 becomes non significant on trend
- **Lecture attendance:** No group effect, trend effect does not change
- **Exam attempts:** No group effect, Item 12 becomes non significant on trend
- **School type:** No change for trend effects, but there are group effects

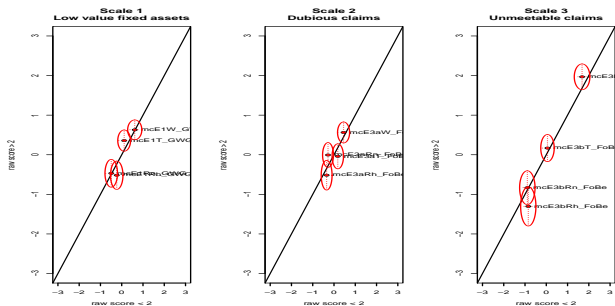
Results: Subject Covariate Effects - II

Differential effects of items at T_2 for different school forms.



Significant are the effects for “HAK” at Item 1, 5, 12 and “other” at Item 3.

Results: Dimensionality and Fairness



- For the MC format unidimensionality and the Rasch Model holds (Andersen and Martin-Loef test).
- For the open format items of scales 2 and 3 are not unidimensional.

- Item 4, 11, 12 are significantly easier at T_2 (as open format)
- All other Items are significantly more difficult at T_2 (as open format)
- The multiple choice format seems to be generally easier
- Covariates have no significant effect on change between the response formats apart from school type
 - “HAK” is a school that teaches accounting, “Other” are tertiary education and technical schools
 - Some items are relatively easier for those two groups at T_2
- MC items for each scale conform to the Rasch model, hence assessment is “fair”
- Caution: Due to the design, we cannot contrast the response formats hence trend and response format effects may be confounded

We suggested to revisit the question with a large scale assessment

- A full assessment of all students in one year (ca. 4500 people)
- Use design that allows contrasting of response format
- Assess the differences individually for each item and for the whole scale

- Fischer, G. H. (1989). An IRT-based model for dichotomous longitudinal data. *Psychometrika*, 52, pp. 599–624.
- Fischer, G. H. (1995). Linear logistic models for change. In G. H. Fischer & I. W. Molenaar (Eds.), *Rasch models: Foundations, Recent developments and Applications* (pp. 157–181). New York: Springer.
- Hatzinger, R. & Rusch, T. (2009). IRT models with relaxed assumptions in eRm: a manual-like instruction. *Psychology Science Quarterly*, 51, pp. 87–120.

Thank you for your Attention

Thomas Rusch

Department of Finance, Accounting and Statistics

Institute for Statistics and Mathematics

email: thomas.rusch@wu.ac.at

URL: <http://statmath.wu.ac.at/~tr>

WU Wirtschaftsuniversität Wien

Augasse 2-6, A-1090 Wien