LASI: modeling students

Diederik M. Roijers

UvA

July 5, 2013

Session questions

- What would we like students, teachers or other stakeholders to know?
- What insights can be gained from analyzing educational data?
- What insights cannot be gained from analyzing educational data?
- In what way can research and educational practice help each other?
- How to do deal with the fact that not the entire learning experience will be captured?

What are we doing?

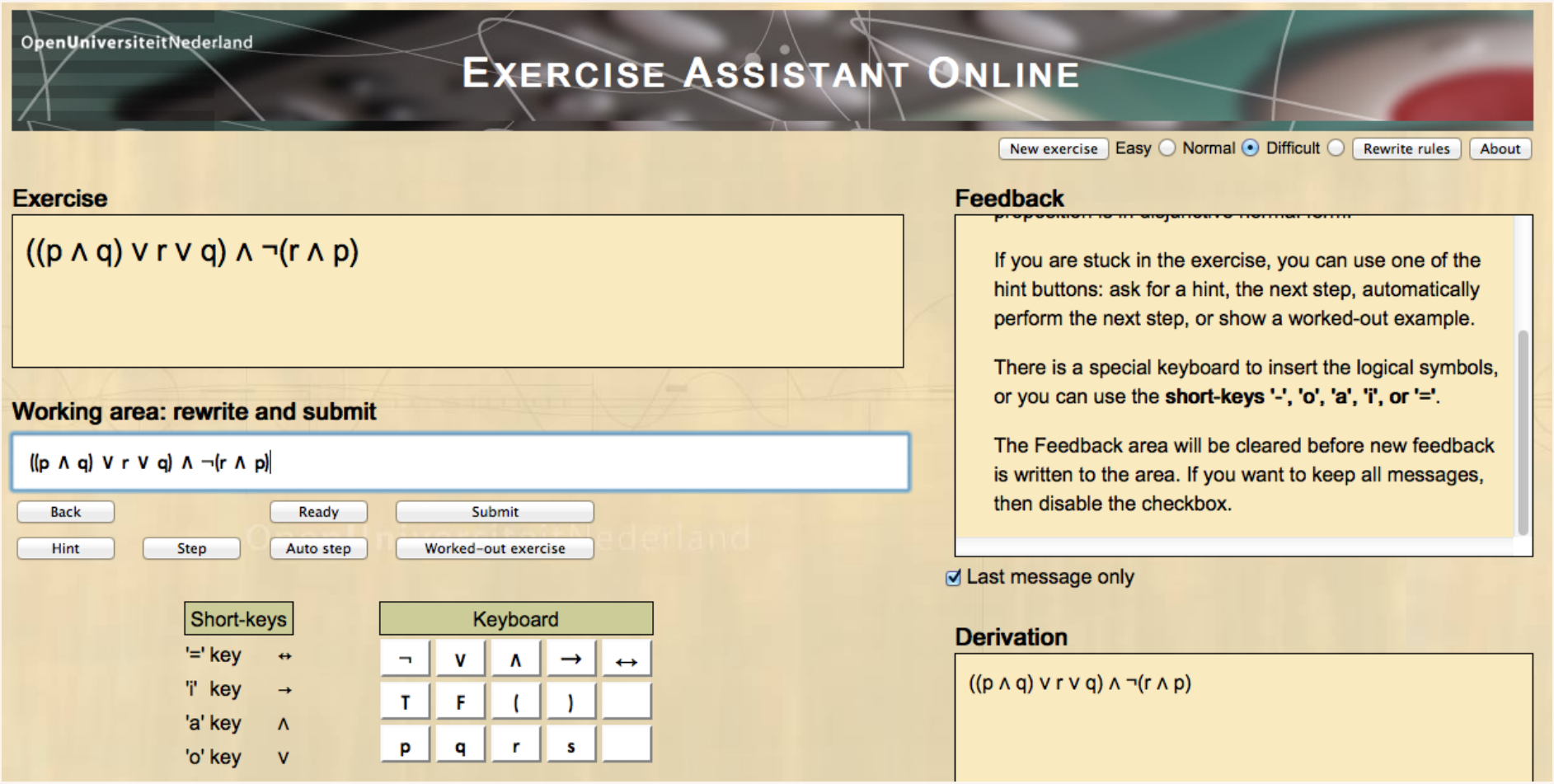
Research

- Student models
  - eIRT (LAK12)
  - Logistic student models
- Adaptive e-tutoring using OER

where

we = [ (Johan Jeuring),  (Ad Feelders),
 (Lieuwe Rekker),  (Sander Latour) ,
 (Maarten van Someren), ]

What data do we want to analyze?



http://ideas.cs.uu.nl/genexas/index.php

What data do we want to analyze?

- 14 students (id)
- 21 rewrite rules (id),
- about 8 instances per rule per student
- Correct (0 or 1)
- Sequential data: 2097 datapoints
- 3 sessions of 2 to 3 hours

What data can we use?

(1,17,0), (1,15,0), (1,3,0), (1,5,0), (1,24,1), (1,5,1), (1,18,0), (1,13,1), (1,11,1), (1,18,0), (1,15,1), (1,3,0), (1,12,1), (1,24,0), (1,8,1), (1,11,1), (1,15,1), (1,12,1), (1,18,0), (1,24,1), (1,7,1), (1,6,0), (1,19,0), (1,10,0), (1,24,1), (1,13,1), (1,12,1), (1,22,0), (1,24,1), (1,16,1), (1,6,1), (1,18,0), (1,17,0), (1,15,1), (1,15,1), (1,6,1), (1,2,1), (1,24,1), (1,18,1), (1,19,0), (1,16,1), (1,15,1), (1,8,1), (1,11,1), (1,24,0), (1,7,0), (1,12,0), (1,3,0), (1,15,1), (1,12,1), (1,17,0), (1,21,0), (1,24,0), (1,3,0), (1,5,0), (1,9,0), (1,17,0), (1,18,0), (1,8,0), (1,11,1), (1,3,0), (1,23,1), (1,6,1), (1,13,1), (1,24,0), (1,15,1), (1,19,0), (1,3,1), (1,20,0), (1,5,1), (1,19,1), (1,24,0), (1,18,1), (1,7,0), (1,10,1), (1,9,0), (1,3,0), (1,9,0), (1,9,1), (1,24,1), (1,21,1), (1,17,0), (1,7,0), (1,9,1), (1,16,1), (1,4,0), (1,24,0), (1,18,0), (1,19,0), (1,17,1), (1,16,0), (1,15,1), (1,24,1), (1,18,0), (1,6,1), (1,13,1), (1,11,1), (1,12,1), (1,24,0), (1,16,1), (1,9,1), (1,12,1), (1,3,1), (1,24,1), (1,19,1), (1,16,1), (1,1,1), (1,17,1), (1,1,1), (1,5,1), (1,1,1), (1,3,1), (1,5,0), (1,24,0), (1,4,0), (1,20,1), (1,20,1), (1,20,0), (1,3,1), (1,3,0), (1,5,1), (1,18,0), (1,16,1), (1,24,0), (1,3,1), (1,18,0), (1,3,1), (1,8,1), (1,12,1), (1,24,0), (1,2,1), (1,18,1), (1,19,1), (1,3,1), (1,24,1), (1,18,1),

What do we want to know?

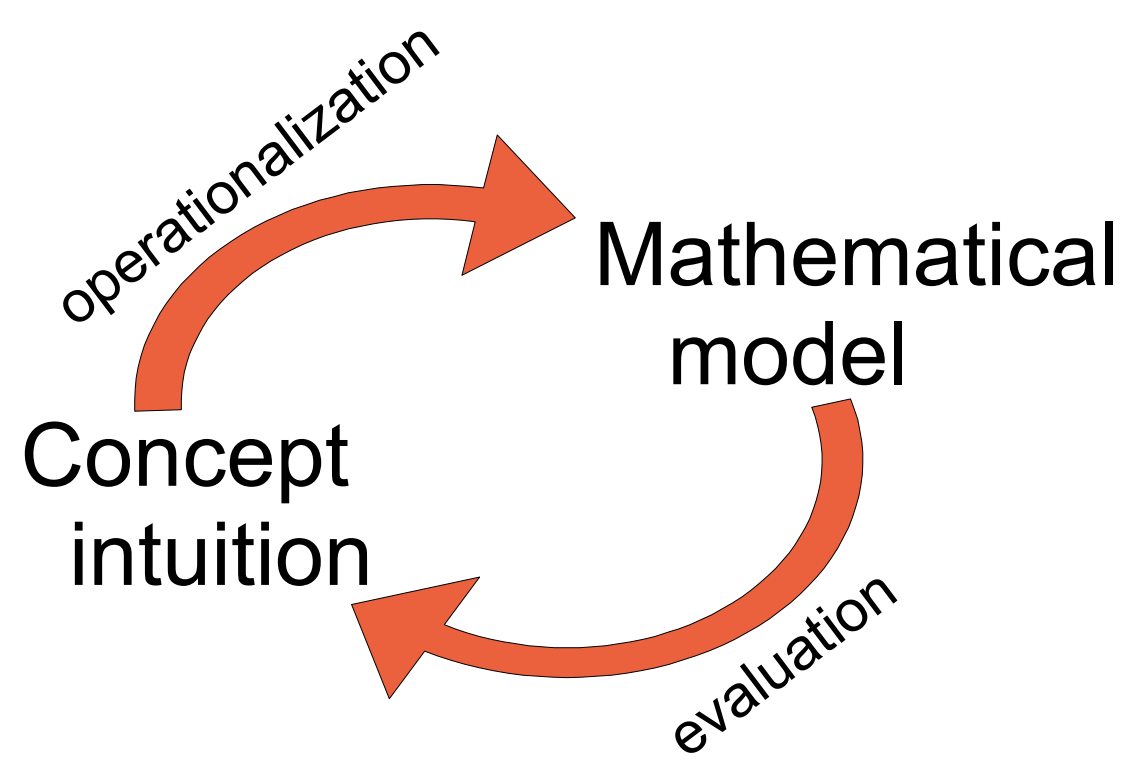
About

- Students
  - Competence
  - Learning
  - Probability of success
  - ...
- Our teaching and tests
  - Difficulty
  - Discriminativity
  - Effectiveness
  - ...

Why

- To improve education
  - Manually
  - Automatically

Modeling


operationalization
Mathematical model
Concept intuition
evaluation

Modeling

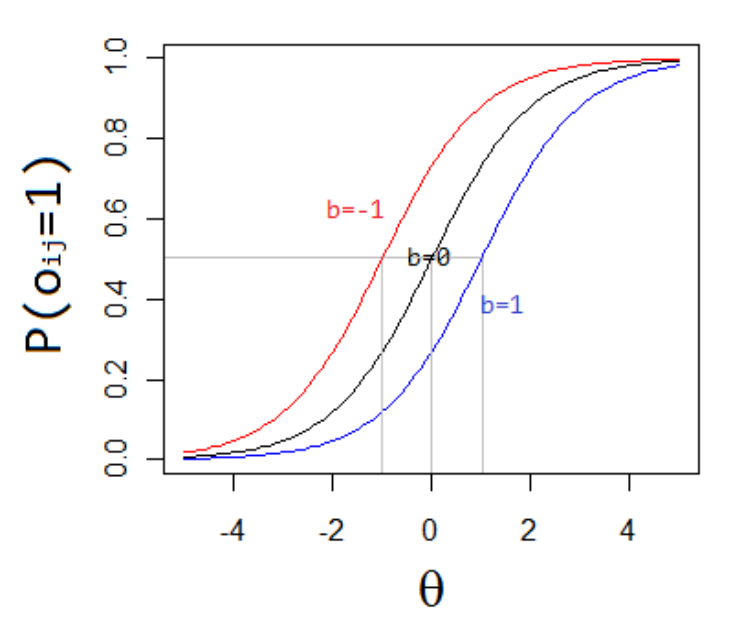
Our vision of a good model

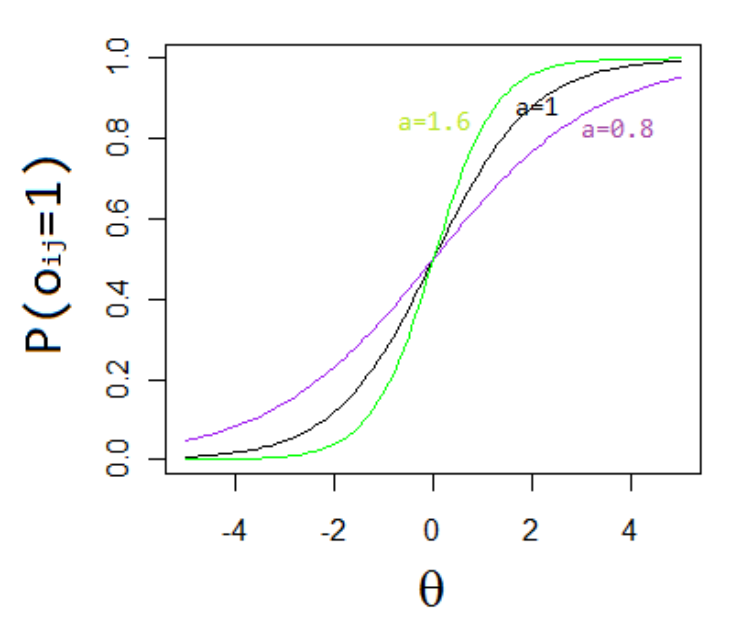
- Makes the right concepts operational
- Corresponds to domain expert intuition
- Is applicable to the data
  - Format
  - Amount
- Is robust
- Is tested
  - Expert evaluation
  - Simulation

Why did we think this would be very doable?

- Very detailed information: every rule application observed
- Reliable starting point: IRT
- Models for much less detailed data exist

Item response theory (IRT): 2PL





Operational definitions of

- difficulty, b
- discriminativity, a
- competence, θ

Logistic function: $\sigma(a(\theta - b))$

$$P(o_{i,j} = 1|\theta_i, b_j, a_j) = \frac{e^{a_j(\theta_i - b_j)}}{1+e^{a_j(\theta_i - b_j)}}$$

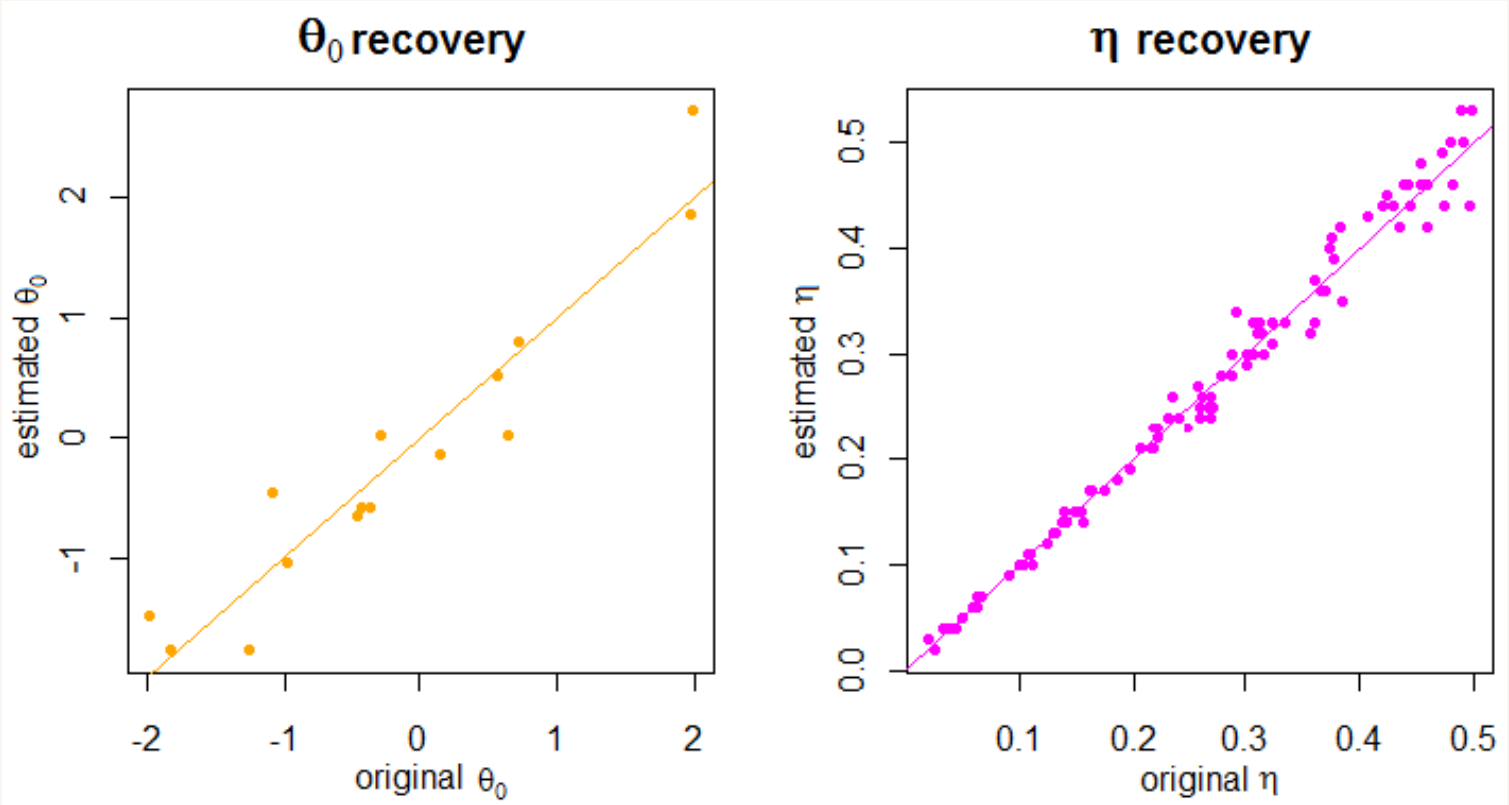
The following probability function describes our model:

$$P(o_{r,s,t} = 1) = \frac{e^{a_r(\theta_{0,s}+\eta_s t_{r,s}-b_r)}}{1 + e^{a_r(\theta_{0,s}+\eta_s t_{r,s}-b_r)}}$$

where r is the rule ID, s the student ID, $\theta_{0,s}$ is the starting competence of the student, η_s the learning speed of the student, and $t_{r,s}$ the number of times student s attempted rule r before.

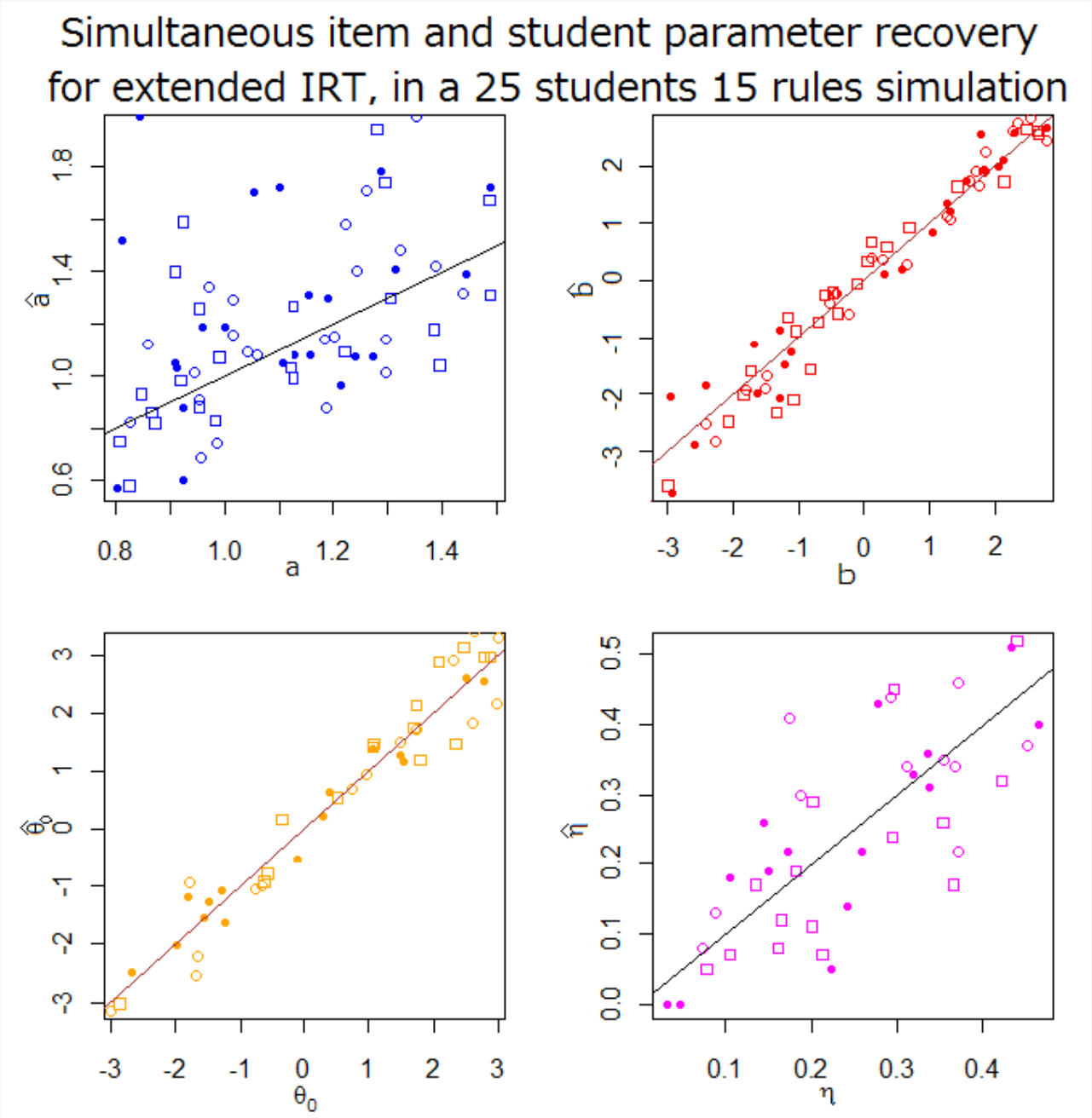
Form chosen using domain experts and viability analysis through simulation.

If we have reliable rule parameter estimates:

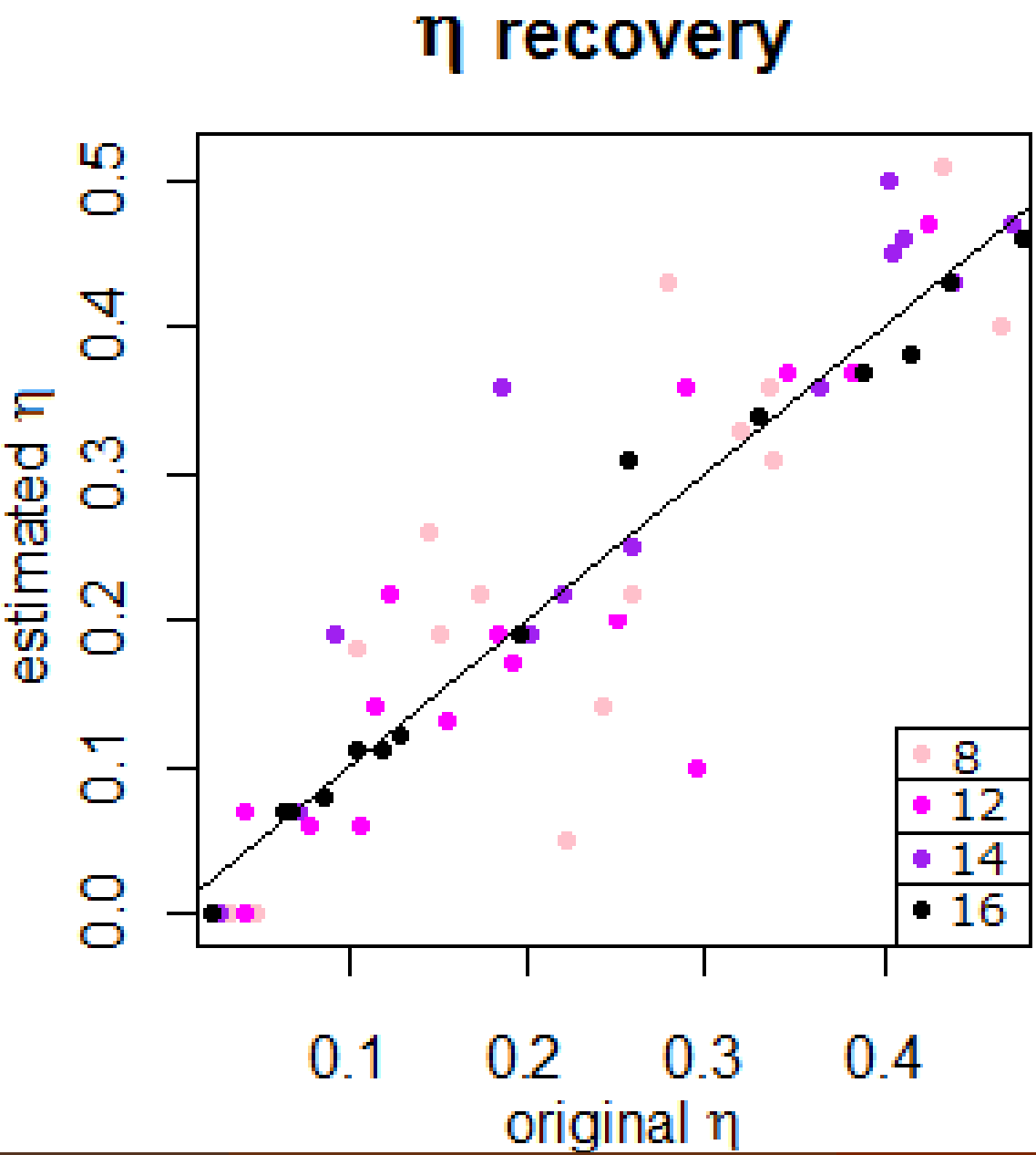


25 rules, 15 students, 8 instances (=data points) per rule.

eIRT: simultaneous rule and student parameter estimation



eIRT: simultaneous rule and student parameter estimation



eIRT: conclusions

- Proven to be able to estimate reliable parameters
- Accuracy 0.8
- Student start competence as expected
- High rank correlation between domain experts and learned parameters for difficulty
- Some surprises (interviews), useful observations

- But, we need more data for η and a

eIRT: lessons learned

Domain experts are really important

As are participant interviews

Be very careful with reliability in relation to the amount of data

Data need for logistic student models

- eIRT is the smallest model we know in the family of logistic student models
- Other models (LFA, PFA) also contain
  - more than one rule per instance/item
  - different learning speeds for correct and incorrect
- but the datasets are usually not much bigger...

Data need for logistic student models

Current research: Lieuwe Rekker

- Generalized model (of which LFA, PFA and eIRT are special cases)
- Compare reliability of logistic student models as a function of the amount of data

Current research goals

- Know what models we can use reliably, with
  - Different assumptions
  - Different amounts of data

Session questions

- What would we like students, teachers or other stakeholders to know?
- What insights can be gained from analyzing educational data?
- What insights cannot be gained from analyzing educational data?
- In what way can research and educational practice help each other?
- How to do deal with the fact that not the entire learning experience will be captured?