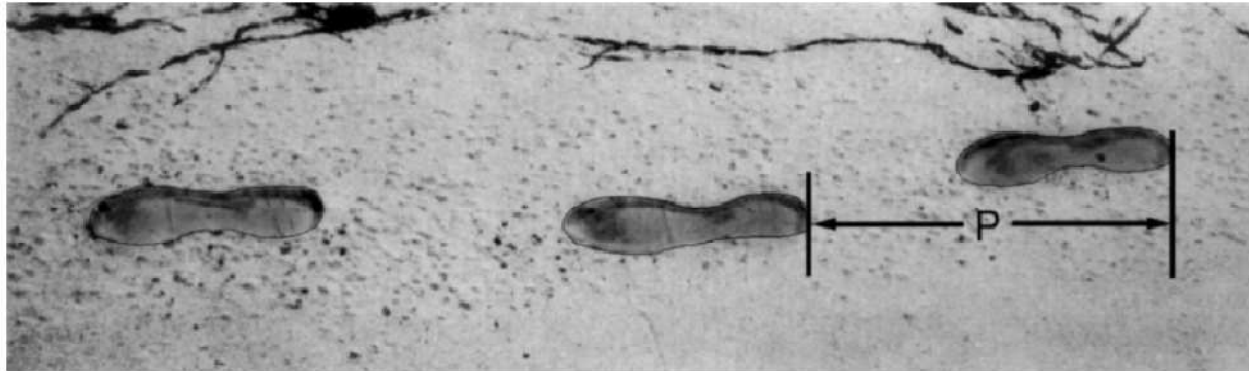


Item Dependency in PISA

PISA Research Conference
Kiel, 14-16 September 2009
C. Monseur

Embedded Structure of PISA Items



The picture shows the footprints of a man walking. The pacelength P is the distance between the rear of two consecutive footprints.

For men, the formula, $\frac{n}{P} = 140$, gives an approximate relationship between n and P where,

n = number of steps per minute, and

P = pacelength in metres

WALKING QUESTION 1

If the formula applies to Heiko's walking and Heiko takes 70 steps per minute. What is Heiko's pacelength?

WALKING QUESTION 3

Bernard knows his pacelength is 0.80 metres. The formula applies to Bernard's walking. Calculate Bernard's walking speed in metres per minute and in kilometres per hour.

Embedded Structure of PISA Items

<i>Cycle</i>	<i>Domain</i>	<i>Number of Units</i>	<i>Number of Items</i>	<i>Average number of Items per Unit</i>
<i>2000</i>	Mathematics	16	31	1.9
	Reading	37	129	3.5
	Science	14	34	2.4
<i>2003</i>	Mathematics	53	84	1.6
	Problem solving	10	19	1.9
	Reading	8	28	3.5
	Science	13	34	2.6
<i>2006</i>	Mathematics	32	48	1.5
	Reading	8	28	3.5
	Science	36	103	2.9

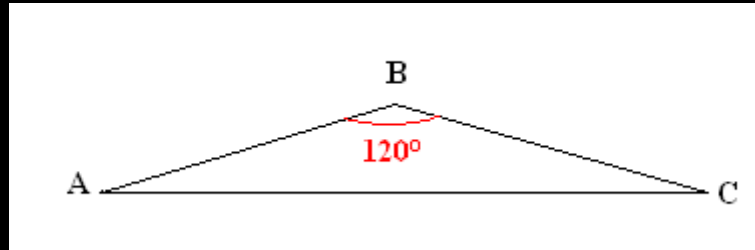
Item Dependency

- Cognitive data scaled with IRT Rasch model
 - Assumption of item independence
 - According to Zenesky et al., *The concept of LID is best understood within the framework of item response theory (IRT). The most popular IRT models specify a single latent trait to account for all statistical dependencies among test items as well as differences among test takers. It is this underlying trait, typically denoted θ , that distinguishes items with respect to difficulty, and distinguishes test takers with respect to proficiency. The probability that a test taker will provide a specific response to an item is a function of the test taker's location on θ and one or more parameters (depending on the IRT model chosen) describing the relationship of the item to θ . Because IRT models are probabilistic, independence must be assumed, conditional on θ , between responses to any pair of items. This conditional independence is called local item independence. [...] When local item dependency is present on a test, inaccurate estimation of item parameters, test statistics, and examinee proficiency may result.*

Item Dependency

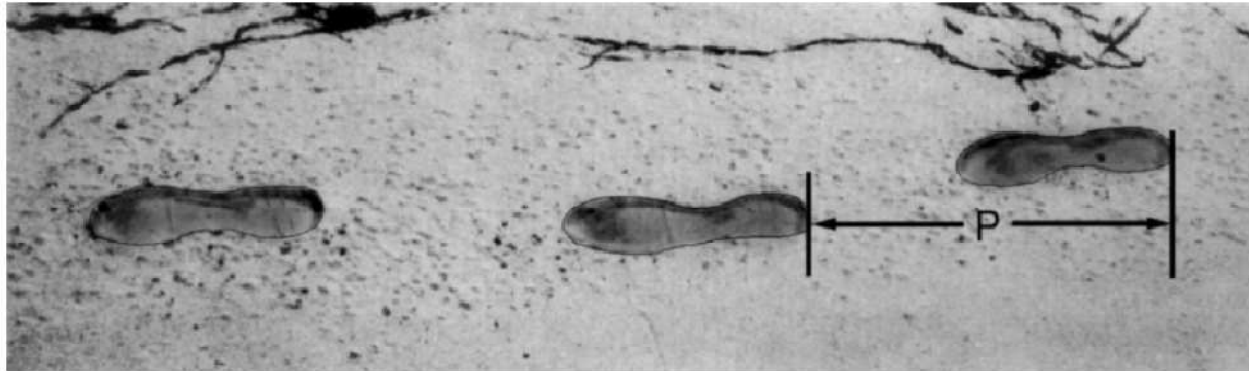
- Probability of succeeding an item is not affected by the success on another item
- Two implemented ways of checking Local Item Dependency
 - Drop of Cronbach α with a unit scaling
 - If no LID, then item Cronbach α and unit Cronbach α are equal
 - Correlation between residual
 - If no LID, correlation between residuals should be equal to 0

Item Dependency



- The triangle ABC is isosceles.
 - 1. What is the measure of angle CAB?
 - 2. Transform this measure in radian?

Embedded Structure of PISA Items



The picture shows the footprints of a man walking. The pacelength P is the distance between the rear of two consecutive footprints.

For men, the formula, $\frac{n}{P} = 140$, gives an approximate relationship between n and P where,

n = number of steps per minute, and

P = pacelength in metres

WALKING QUESTION 1

If the formula applies to Heiko's walking and Heiko takes 70 steps per minute. What is Heiko's pacelength?

WALKING QUESTION 3

Bernard knows his pacelength is 0.80 metres. The formula applies to Bernard's walking. Calculate Bernard's walking speed in metres per minute and in kilometres per hour.

Item Dependency

	Unit 1			Unit 2				Unit 1	Unit 2
	Q1	Q2	Q3	Q1	Q2	Q3	Q4		
<i>Student 1</i>	1	0	2	9	1	9	R	3	1
<i>Student 2</i>	0	1	1	0	1	1	1	2	3
<i>Student 3</i>	1	1	2	1	1	0	9	4	2
<i>Student 4</i>	1	1	0	1	1	1	1	2	4
<i>Student 5</i>	0	1	1	1	0	1	0	2	2

- Non reached items
 - PISA: missing for calibration, incorrect answer for proficiency estimate
 - Here: incorrect answer for both steps
- Suppression of whole unit if one item was deleted at the national level

Drop in Cronbrach α

Decrease in the reliability estimate per booklet and per domain in PISA 2000

Booklet	Mathematics	Reading	Science
1	0.08	0.03	
2		0.07	0.05
3	0.08	0.04	
4		0.04	0.06
5	0.10	0.03	
6		0.04	0.05
7		0.05	
8	0.06	0.06	0.09
9	0.09	0.04	0.02

Drop in Cronbrach α

Decrease in the reliability estimate per booklet and per domain in PISA 2003

Booklet	Mathematics	Reading	Science	Problem
1	0.02	0.08		
2	0.03	0.04		
3	0.03			0.03
4	0.03			0.03
5	0.03		0.05	
6	0.03		0.05	
7	0.03	0.09	0.05	
8	0.09	0.03	0.02	
9	0.09	0.08	0.05	0.03
10	0.10	0.03		0.03
11	0.03	0.02		0.02
12	0.04		0.06	0.01
13	0.03		0.05	0.02

Correlation between residuals



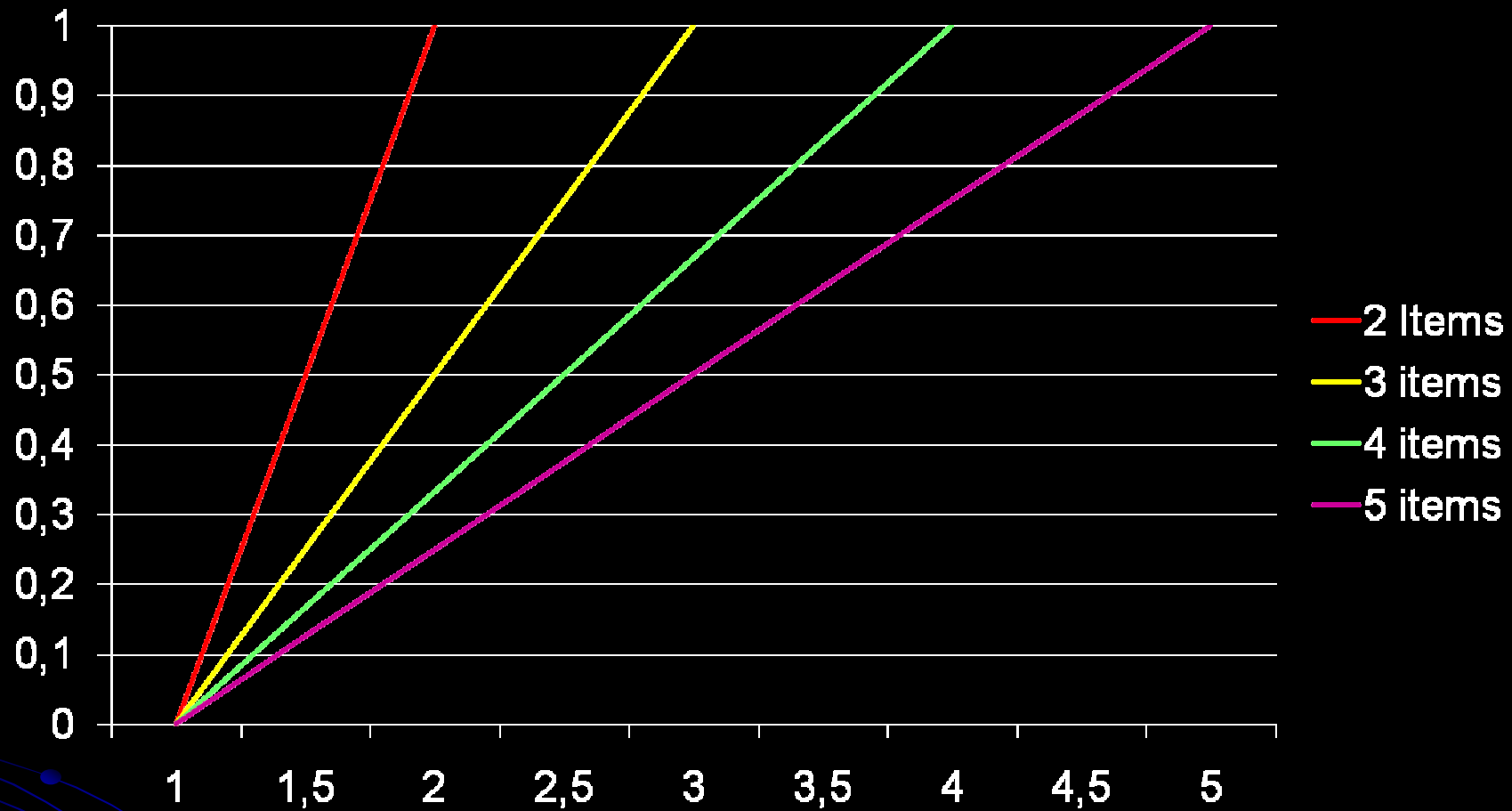
- Computation of the residual
 - Difference between the expected score and the observed score per item and per student;
 - Principal component analysis on residual per unit
 - Computation of an indice based on the first eigenvalue

Correlation between residuals

- Index of dependency that varies between 0 and 1

$$\text{Coef}_{dep} = \frac{\frac{\sigma^2}{nb_{items}} \frac{1}{nb_{items}}}{1 - \frac{1}{nb_{items}}} \frac{\sigma^2 - 1}{nb_{items} - 1}$$

Correlation between residuals

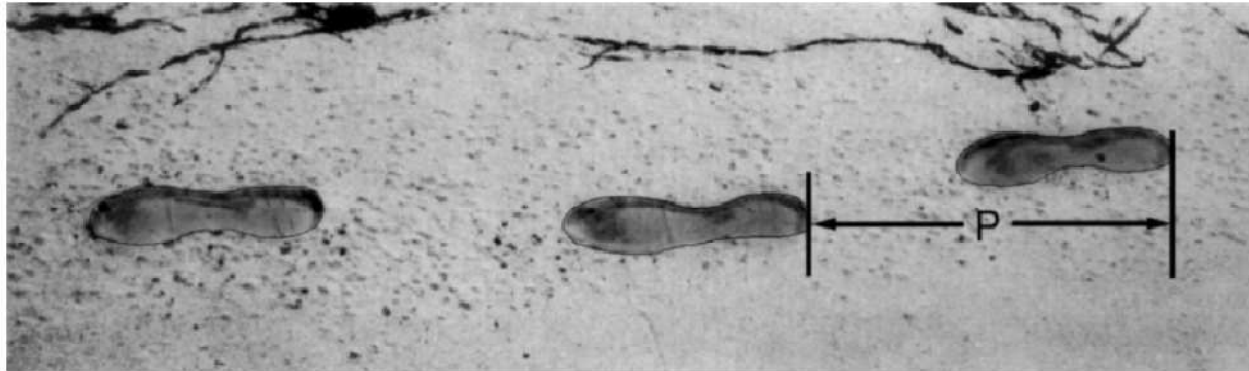


Correlation between residuals

- **summary statistics of the dependency coefficient in mathematics PISA 2000**

UNIT	N	MIN	MAX	MEAN	STD
M124	36	0.00	0.38	0.21	0.10
M136	35	0.06	0.25	0.14	0.03
M144	36	0.06	0.19	0.10	0.03
M150	36	0.01	0.24	0.10	0.05
M037	36	0.01	0.23	0.07	0.06
M155	32	0.02	0.10	0.05	0.02
M159	36	0.02	0.09	0.05	0.02

PISA 2000 M124 Items



The picture shows the footprints of a man walking. The pacelength P is the distance between the rear of two consecutive footprints.

For men, the formula, $\frac{n}{P} = 140$, gives an approximate relationship between n and P where,

n = number of steps per minute, and

P = pacelength in metres

WALKING QUESTION 1

If the formula applies to Heiko's walking and Heiko takes 70 steps per minute. What is Heiko's pacelength?

WALKING QUESTION 3

Bernard knows his pacelength is 0.80 metres. The formula applies to Bernard's walking. Calculate Bernard's walking speed in metres per minute and in kilometres per hour.

Correlation between residuals

- **summary statistics of the dependency coefficient in reading PISA 2000**

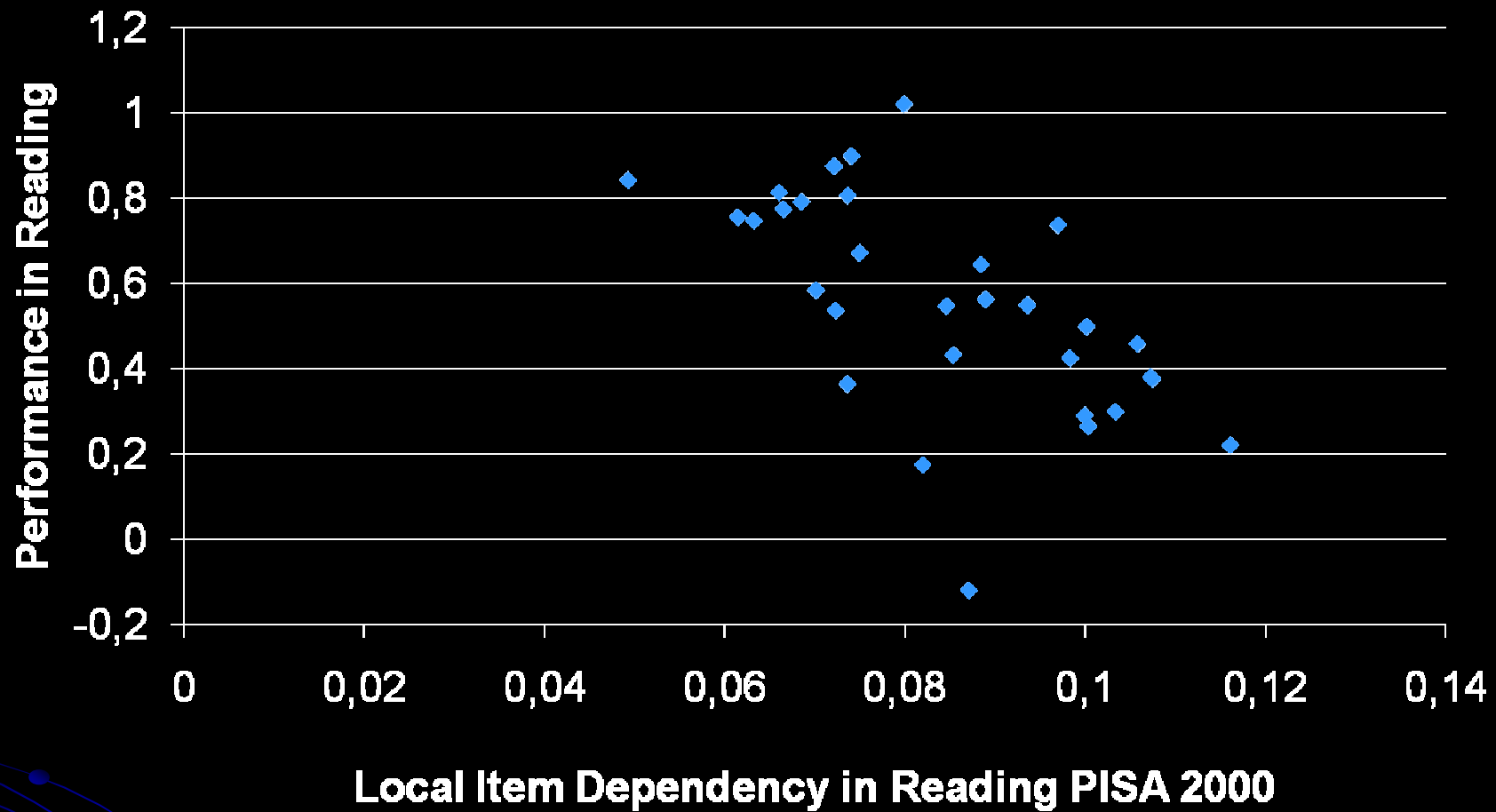
UNIT	N	MIN	MAX	MEAN	STD
R067	36	0.02	0.37	0.22	0.06
R219	35	0.09	0.32	0.20	0.06
R076	31	0.08	0.30	0.19	0.05
R220	36	0.07	0.29	0.16	0.05
R100	34	0.04	0.27	0.13	0.06
R104	36	0.04	0.22	0.13	0.05
R239	34	0.01	0.29	0.13	0.08

Correlation between residuals


- **summary statistics of the dependency coefficient in science
PISA 2000**

UNIT	N	MIN	MAX	MEAN	STD
S114	36	0.05	0.20	0.11	0.04
S195	35	0.03	0.18	0.07	0.03
S268	31	0.02	0.15	0.06	0.03
S129	36	0.00	0.26	0.05	0.05
S131	36	0.00	0.16	0.05	0.04
S252	36	0.02	0.12	0.05	0.03
S253	36	0.02	0.14	0.05	0.03

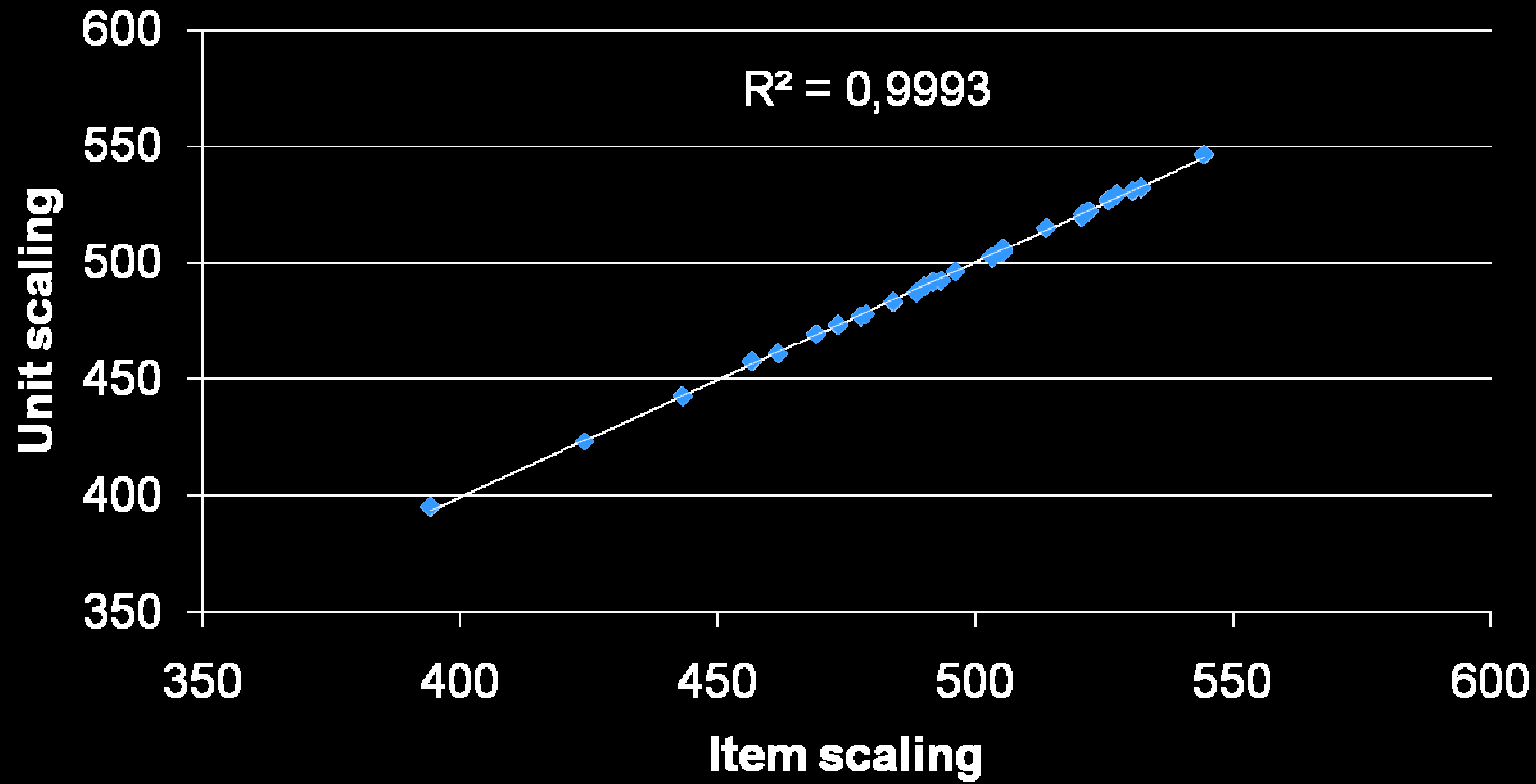
Correlation between residuals



Consequences on population estimates

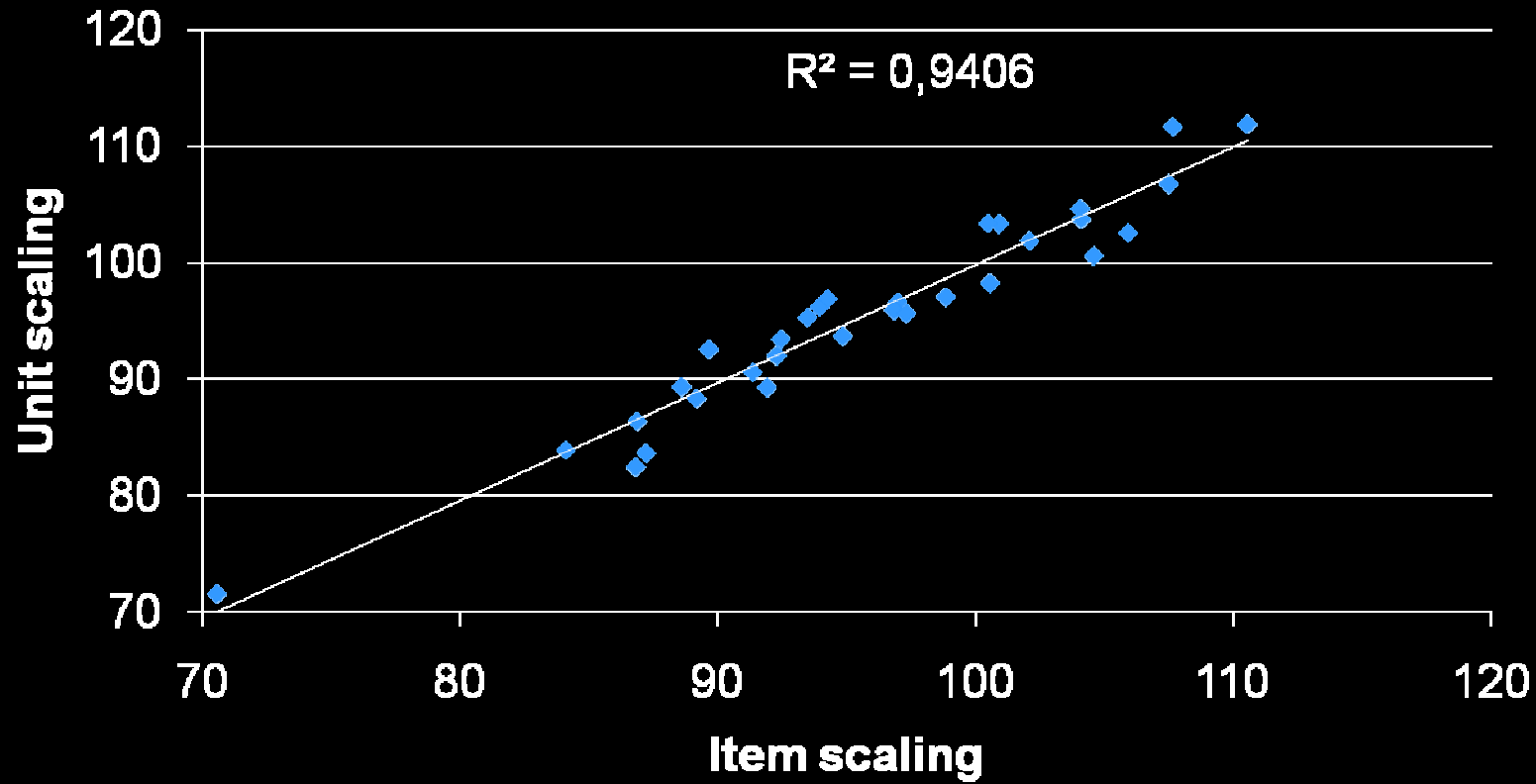
- The assumption of Local Item Independence is violated
 - Consequences on performance indicators
 - Mean estimates in performance
 - Consequences on equity indicators
 - STD in performance
 - Student estimates: WLE, not PVs
- 

Mean estimates in reading (PISA2000)



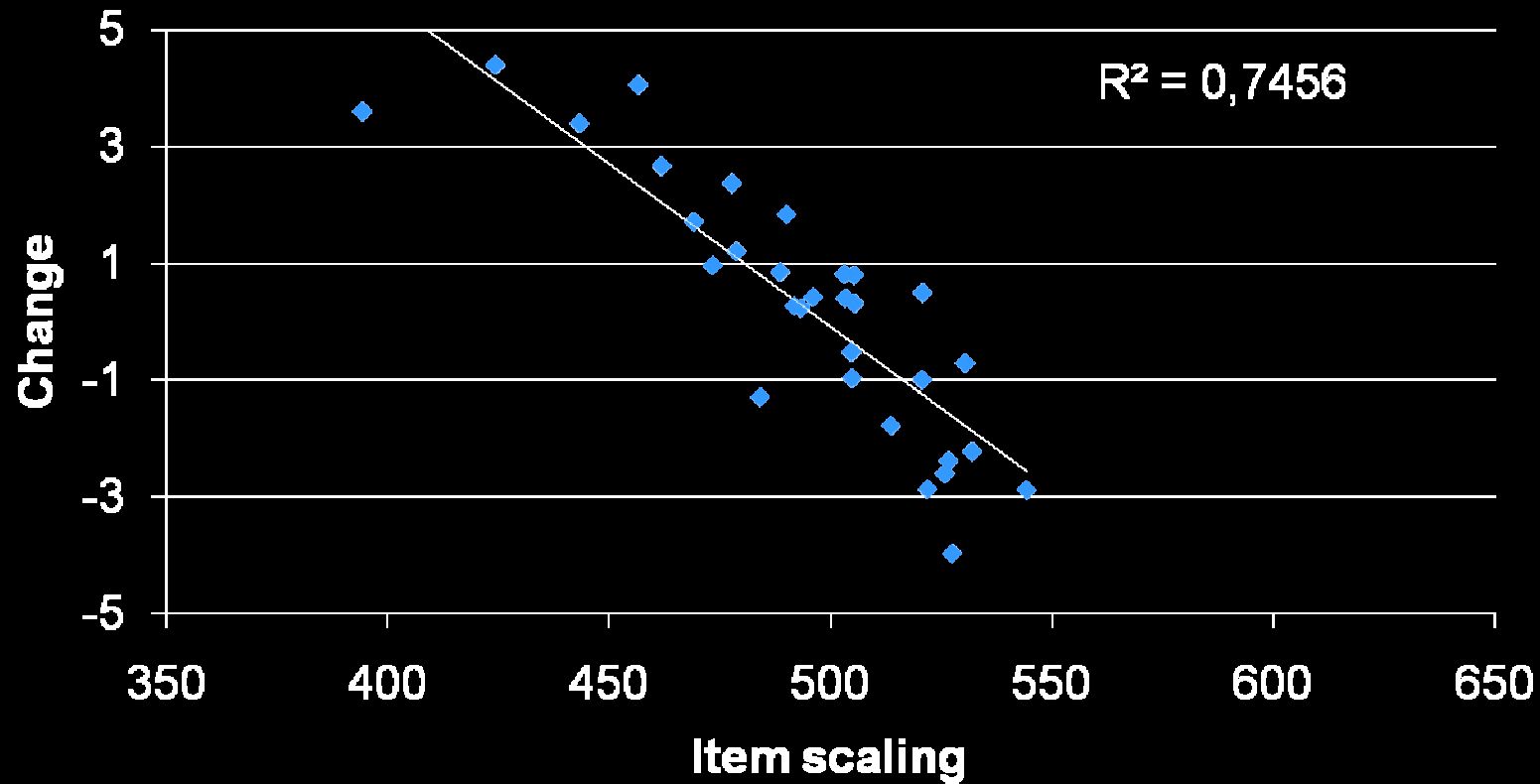
- Correlation > 0,99 in mathematics and in science

STD estimates in reading (PISA2000)



- Correlation > 0,97 in mathematics and in science

Change in STD estimates in reading (PISA2000)



- Correlation = -0,70 in math and -0,30 in science

Change in STD estimates in reading (PISA2000)

	Perfomance	Item STD	Unit STD	Change
NSL	527,5	107,7	111,6	-4,0
FIN	544,4	89,7	92,5	-2,9
GBR	522,0	100,5	103,4	-2,9
IRL	525,9	94,3	96,9	-2,6
AUS	526,8	100,9	103,3	-2,4
CAN	532,1	94,0	96,2	-2,2
POL	477,8	100,6	98,2	2,4
RUS	461,8	91,9	89,3	2,7
LUX	443,3	105,9	102,6	3,4
BRA	394,2	87,2	83,6	3,6
LVA	456,6	104,6	100,6	4,0
MEX	424,2	86,8	82,4	4,4

Conclusions

- The structure of the PISA items violates more or less the Local Item Independence of the IRT models
 - This LID does change
 - Proficiency estimates (% per level)
 - Deviation estimates
 - Further, the change in STD correlates with achievement

Conclusions

- LID should be used during
 - Data cleaning
 - Item revision after the FT
 - Item selection for the MS
- This example illustrates the impact of methodological choice on population estimate
- It also emphasises the relative character of indicators