Modelling students progression in secondary education with the combination of separate cohorts of repeated measurements

Maria Pampaka & Patricio Troncoso
The University of Manchester, UK
maria.pampaka@manchester.ac.uk

Overview

- Introducing the (educational) problem
- Introducing the project

- The analytical/methodological framework
  - brief overview of instruments
  - brief overview of measure construction

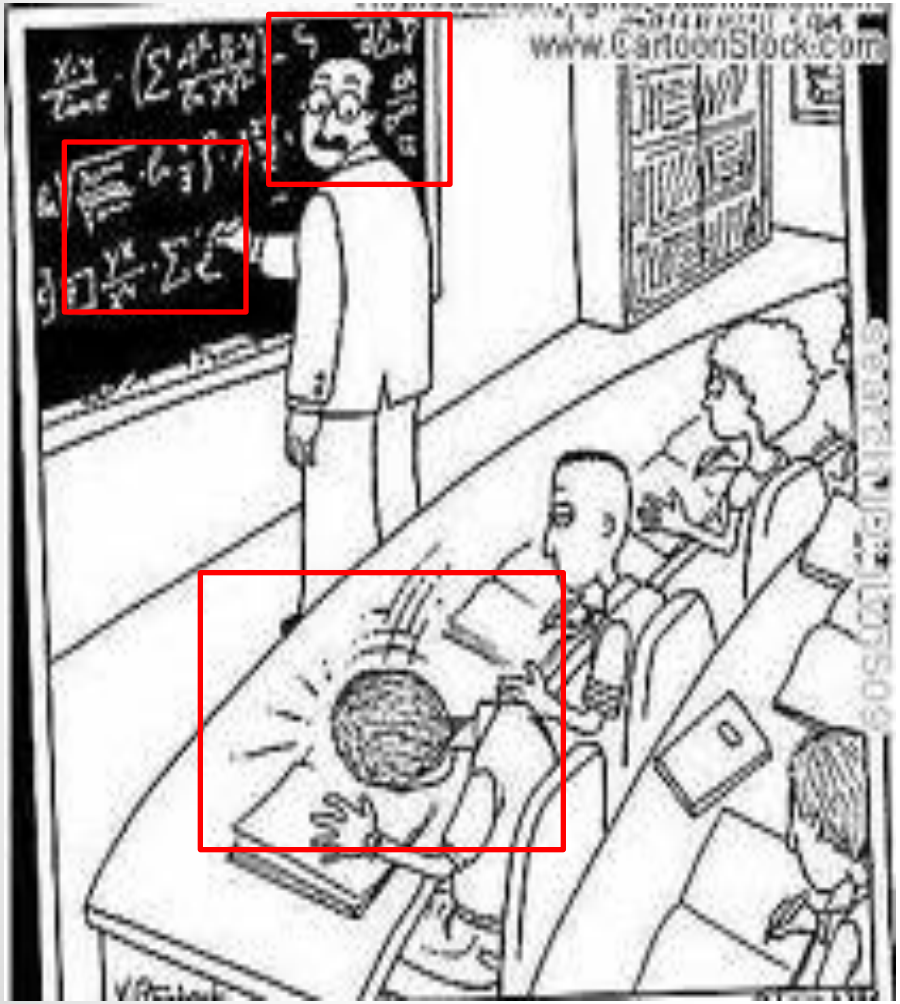
- Statistical modelling of repeated measures (of dispositions)
  - Some descriptive results
  - Analytical challenges
  - Models of change over time

- Concluding Points / Further challenges and analysis

The problem…

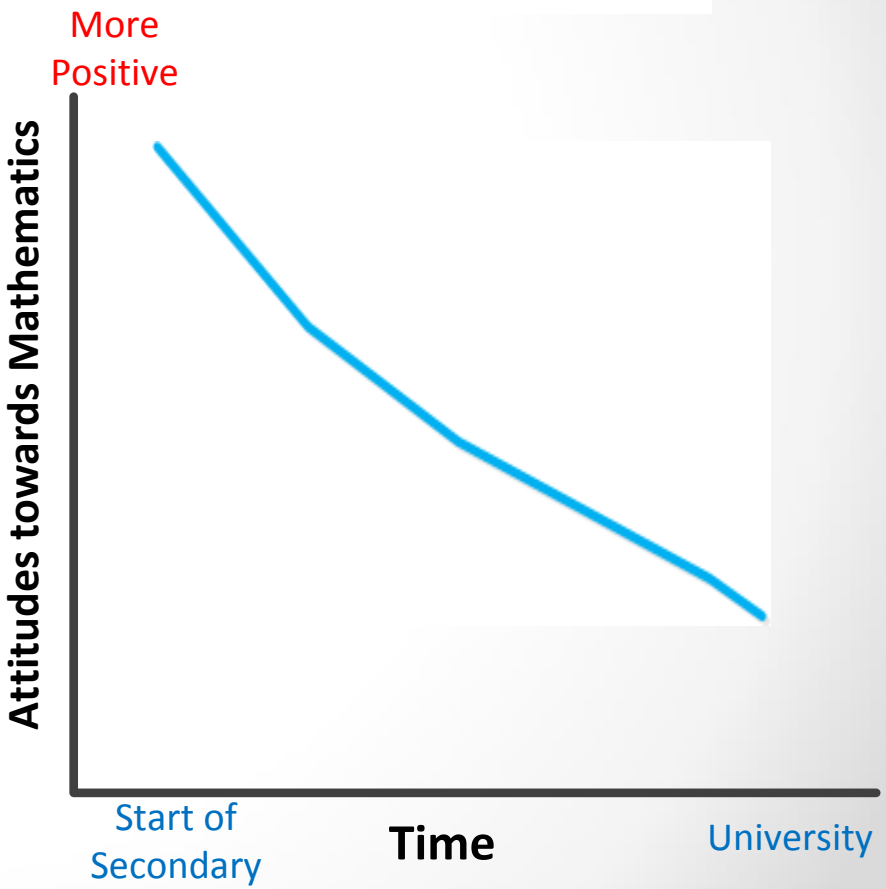
Declining students' mathematics dispositions/attitudes

The main actors

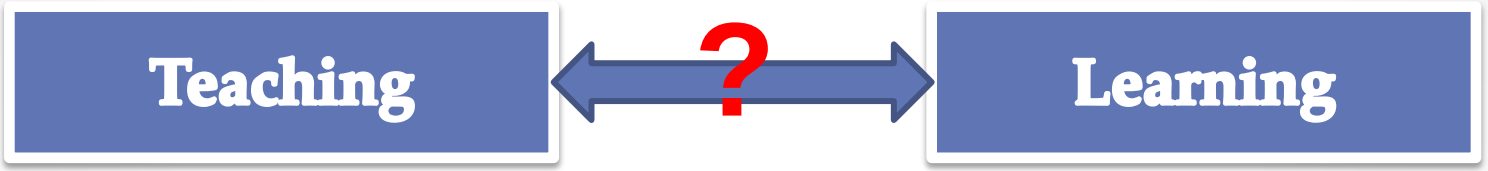


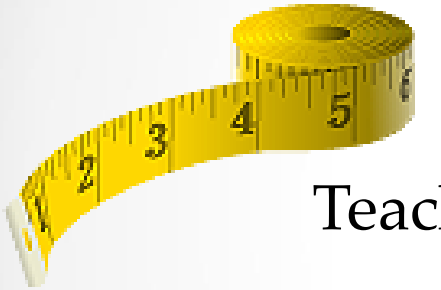
Recent evidence

…



A research problem / question…





Teaching practices

Learning Outcomes and Attitudes

What is the association between teaching styles/practices in mathematics with variables relevant to students' mathematical dispositions /attitudes?

TeLePriSM (Acknowledgements)

Teaching and Learning Practices in Secondary Mathematics

Ongoing ESRC funded study (RES-061-25-0538) in UK (2011-2014) (www.teleprism.com)

"Mathematics teaching and learning in secondary schools: the impact of pedagogical practices on important learning outcomes"

TEAM

Project investigator	Maria Pampaka
Research Associates	Lawrence Wo, Afroditi Kalambouka
Associate Research students	Sophina Qasim, David Swanson, Patricio Troncoso-Ruiz
Mentors	Prof Julian Williams, Prof Ian Plewis

Teleprism: Aims

Aim: To map secondary students' learning outcomes and choices, including dispositions and attitudes, together with the teaching they are exposed to.

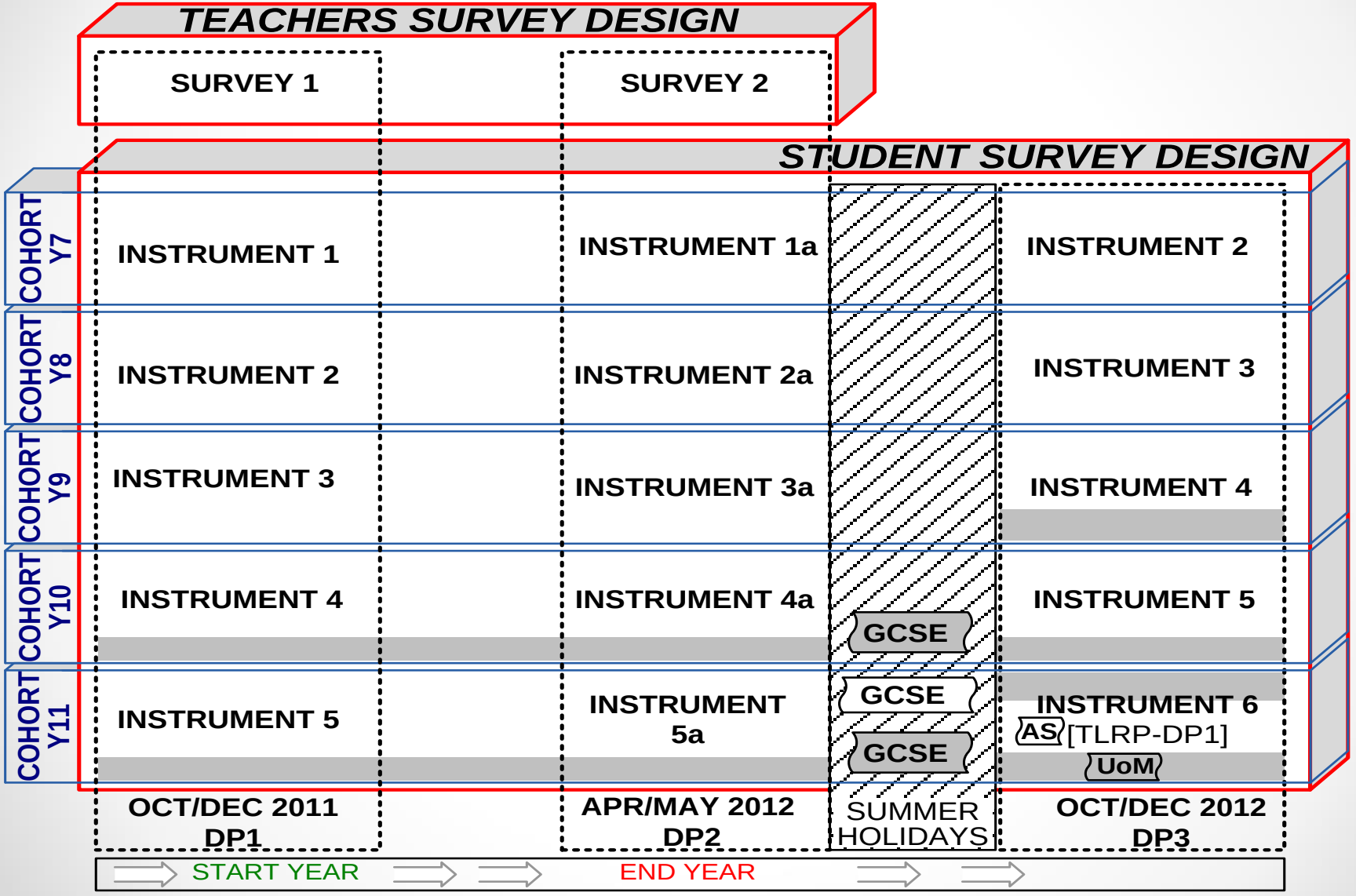
- Surveys for students from Years 7 to 11 (3 times) and also for their mathematics teacher (twice).

- Case studies in a small number of schools with lesson observations and interviews with students and teachers.

- Note: UK secondary compulsory education

  Year 7 (age 11) to Year 11 (age ~16, GCSE exams)

The Teleprism Survey Design



Participating Schools

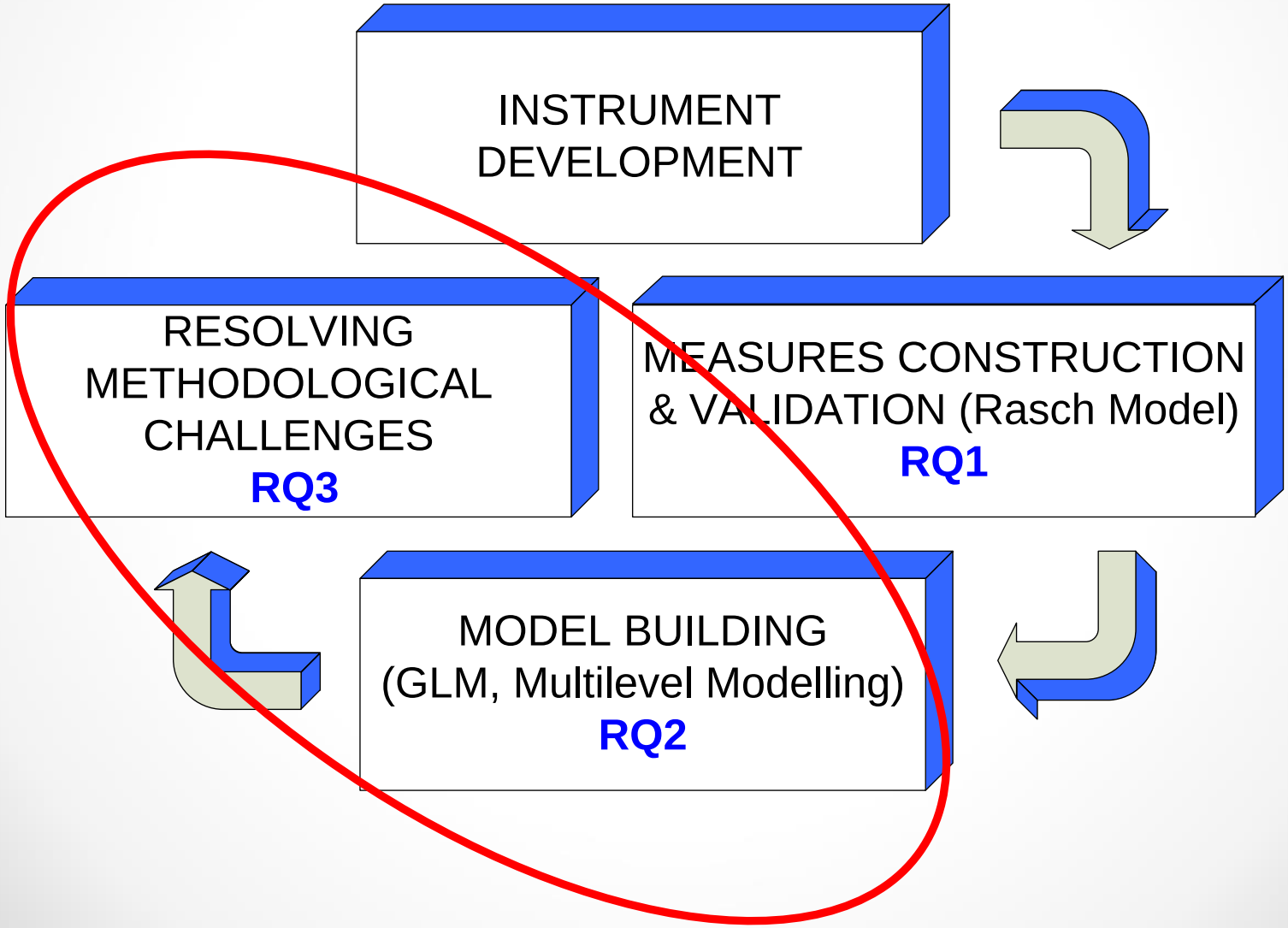
Age range	Boys only	Girls only	Mixed	Total
11-16	0	2	13	15
11-18	1	5	19	25
Total	1	7	32	40

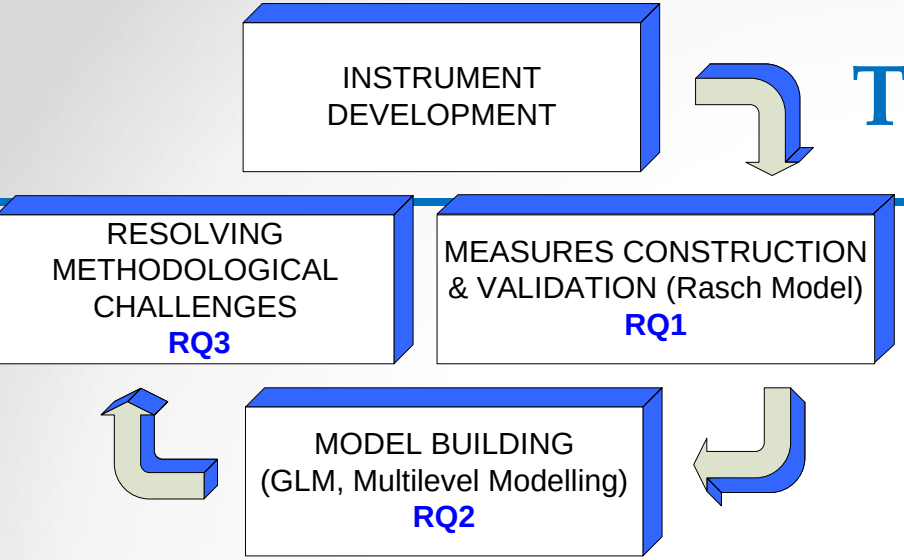
Students @ start

Year 7	3884
Year 8	3025
Year 9	2668
Year 10	2145
Year 11	1794
Total	**13516**

TeLePriSM

The methodological/Analytical Framework





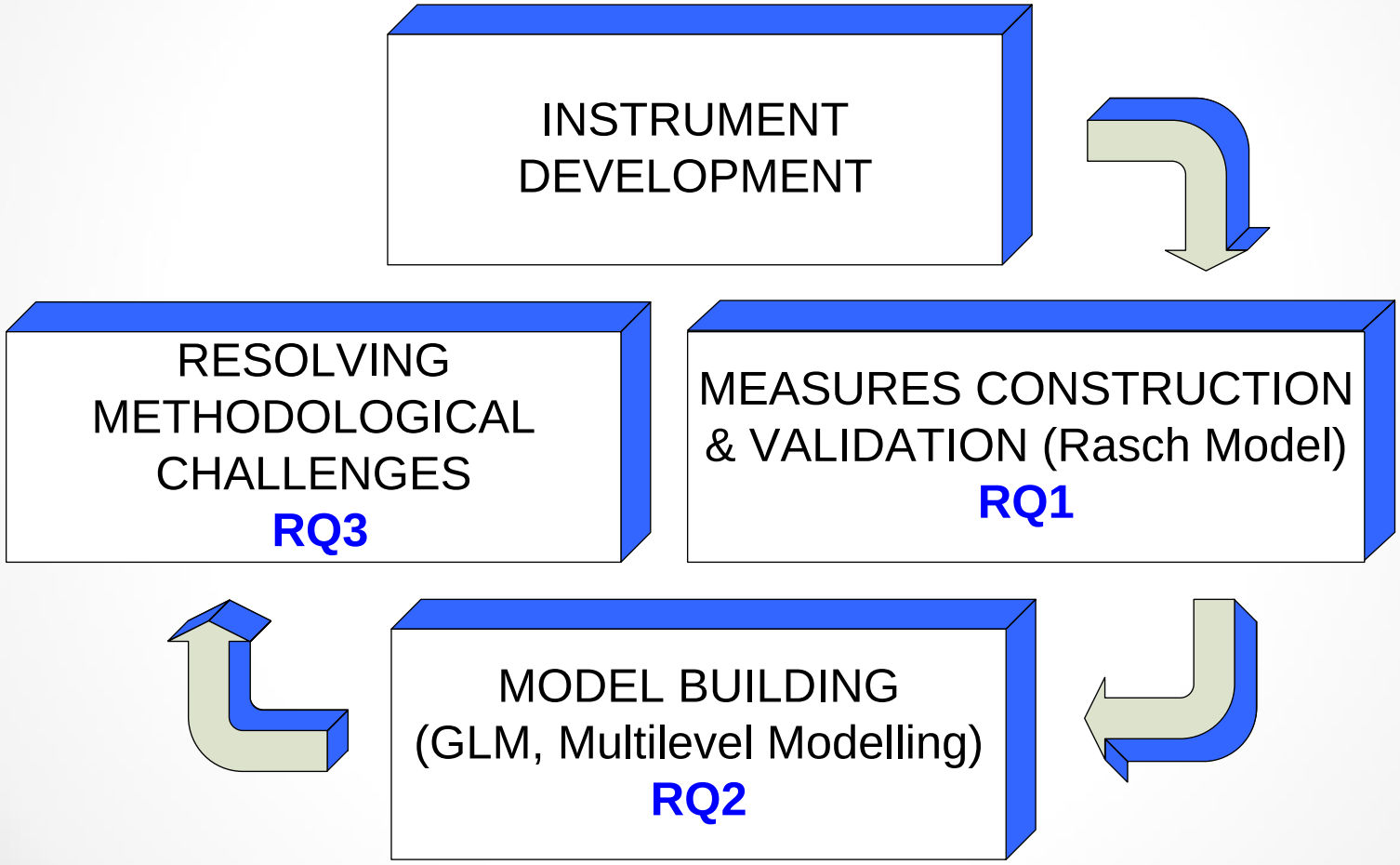
The Research Questions

RQ1: How can we measure (i) teachers' (self-reported) pedagogic practices and (ii) students' dispositions (and other learning outcomes) to study and use mathematics?

How do these measures vary across key subgroups (e.g. year groups), background variables (e.g. class, ethnicity, gender) and institutional types (schools)?

- **RQ2:** How do background and process variables (e.g. programme type) and pedagogy predict students' learning dispositions, outcomes and decisions from Y7 to Y11?
- **RQ3:** How can cross-sectional and longitudinal models be combined in the context of hierarchical data structures and missing data?

The methodological/Analytical Framework

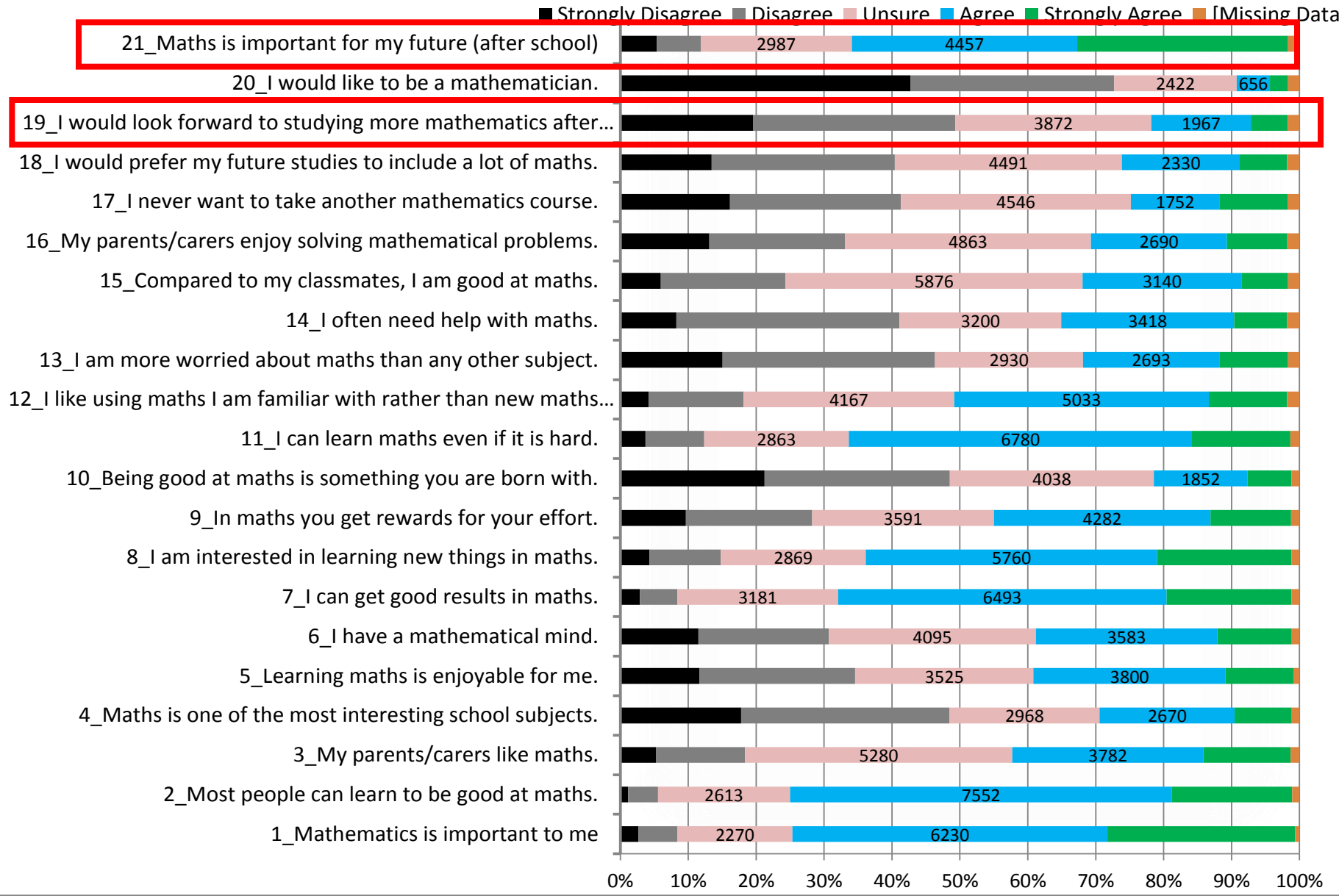


INSTRUMENT DEVELOPMENT

The Questionnaire

- About yourself and your school
  - Background information
  - Class and Teacher identifiers
  - Parental support/involvement
- Your feelings about mathematics (Maths Attitudes)
- Aspirations and intentions for after High School
- How maths is taught (Perceptions of teaching)
- Confidence in maths tasks (Maths Self-efficacy)

Example: Maths Attitudes



Example: Maths Self-Efficacy

5. How confident are you to **calculate the range of a set of numbers** such as:

A rugby team played 7 games.

Here is the number of points they scored in each game.

3 5 8 9 12 12 16

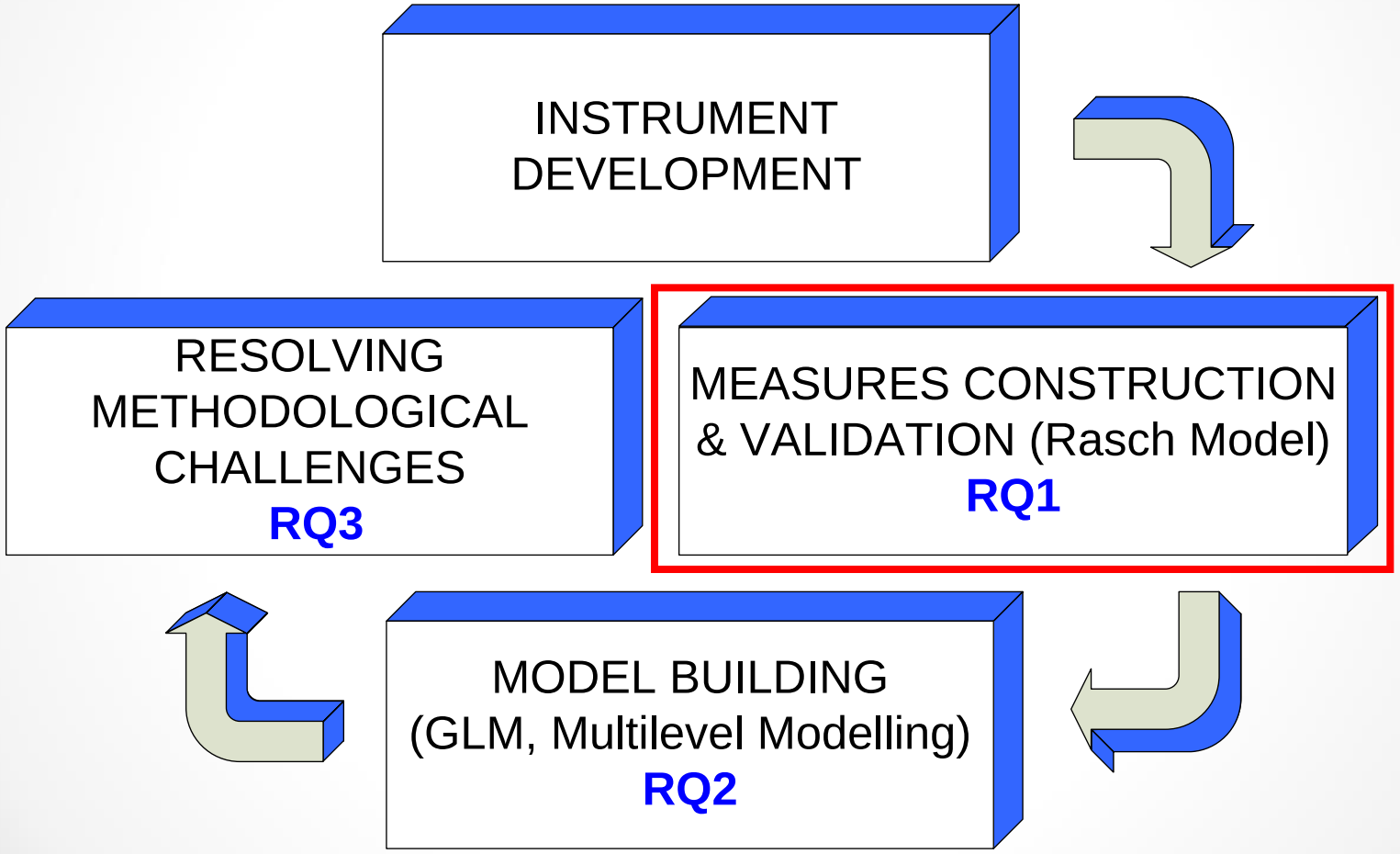
(a) Work out the range.

.....

Not confident at all **Not very confident** **Fairly confident** **Very confident**

"In this section, we are asking you to say how confident you would be at using mathematics to solve different problems. We don't ask you to actually solve the problems."

The methodological/Analytical Framework



MEASURES CONSTRUCTION & VALIDATION (Rasch Model)
RQ1

Measurement

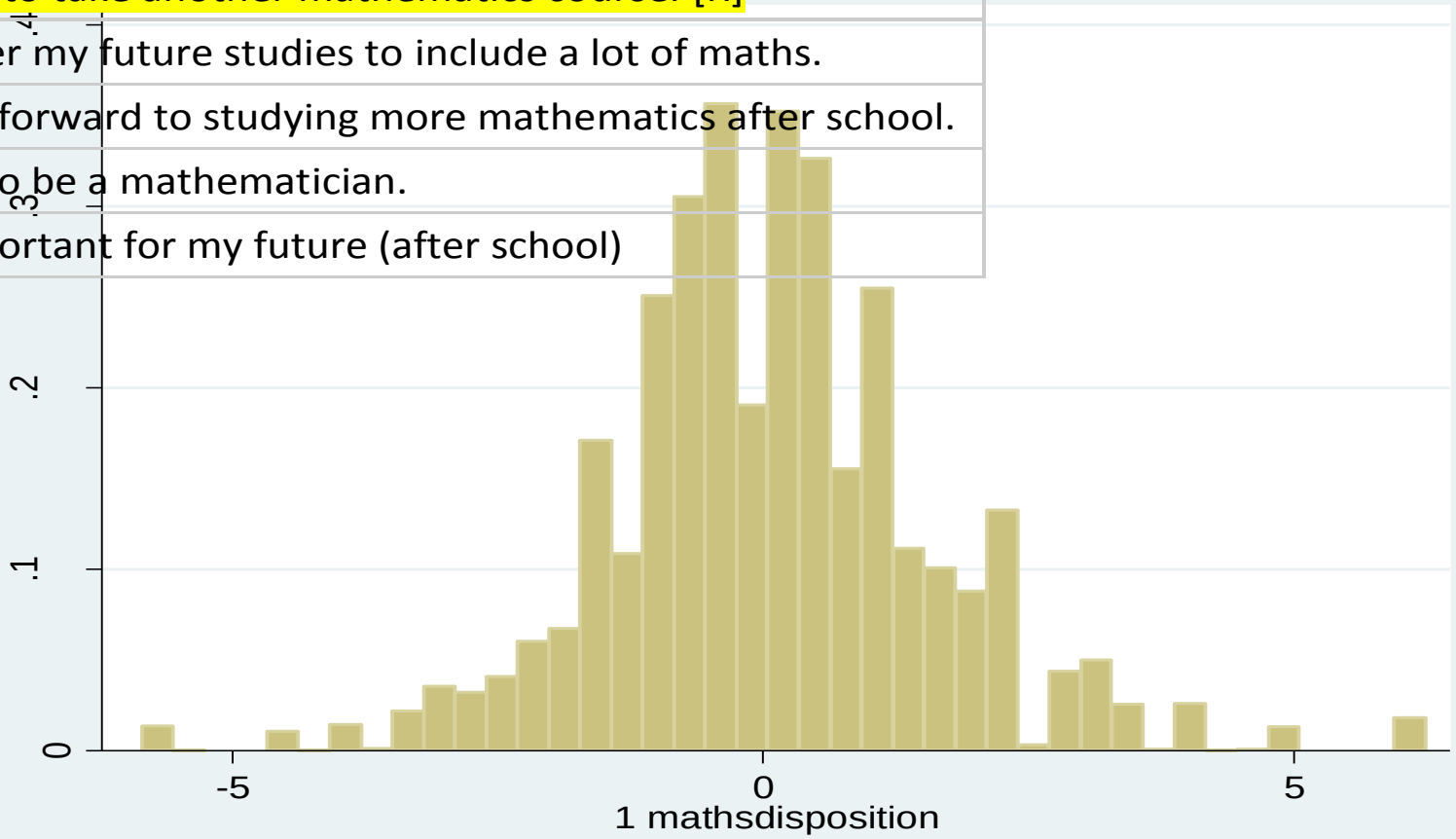
- 'Theoretically': Rasch Analysis (IRT)
  - Partial Credit Model
  - Rating Scale Model

- 'In practice' – the tools: Winsteps software

- Interpreting Results:
  - Item Fit Statistics (to ensure unidimensional measures)
  - Differential Item Functioning for 'subject' groups
  - Person-Item maps for hierarchy
  - Differential Item Functioning (DIF)
  - Qualitative checks

Example: A measure of maths disposition

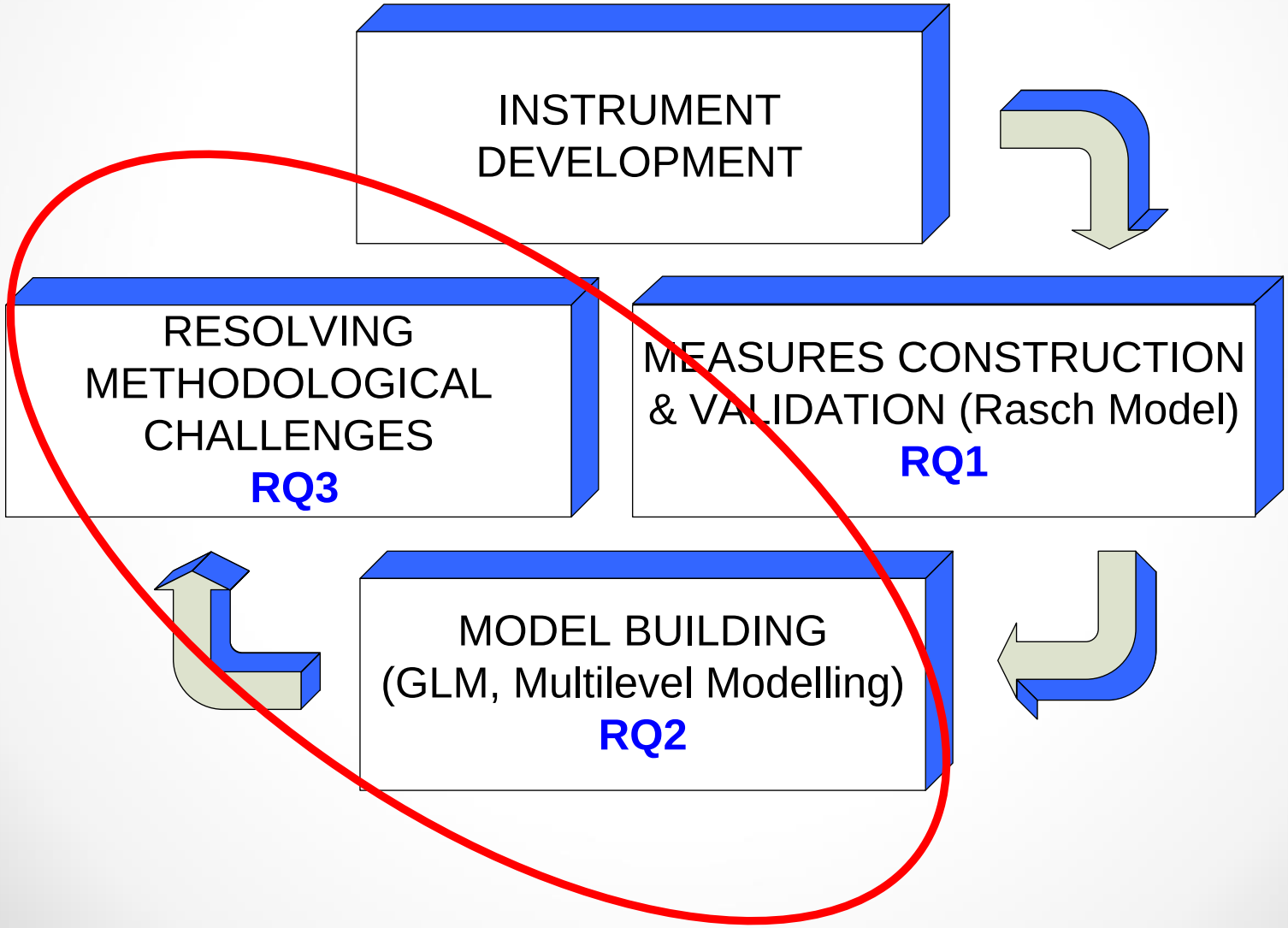
1	Mathematics is important to me.
4	Maths is one of the most interesting school subjects.
5	Learning maths is enjoyable for me.
8	I am interested in learning new things in maths.
17	I never want to take another mathematics course. [R]
18	I would prefer my future studies to include a lot of maths.
19	I would look forward to studying more mathematics after school.
20	I would like to be a mathematician.
21	Maths is important for my future (after school)



Constructed and validated measures

- A measure of 'perceived parental involvement/support'
- Mathematics disposition: (the higher the score the more disposed the student is towards further study or engagement with mathematics)
- Mathematics 'identity': (the higher the score the more positively/strongly the student relates or identifies with mathematics)
- Mathematics Self-efficacy
- Perceptions of teaching:
  - Teaching Variation: the higher the score on this measure the more diverse the maths lessons are (from students' perspective).
  - Transmissionist teaching: the higher the score the more 'traditional' or teacher-centred the practices as reported by the students.

The methodological/Analytical Framework



Subjects and patterns of completion

- Unique cases of students who took part in the study: 18 157
- Unique student ids managed by schools (ethical constraints)

Challenge 0: Matching students responses across DPs to enable longitudinal analysis

Freq.	Percent	Cum.	Pattern
5830	32.11	32.11	Only DP1
3629	19.99	52.10	**All DPs**
2992	16.48	68.57	DP1 and DP2
2453	13.51	82.08	Only DP3
1298	7.15	89.23	Only DP2
1179	6.49	95.73	DP1 and DP3
776	4.27	100.00	DP2 and DP3
18157	100.00		

Sample per Year group/cohort

Year Group @	DP1	DP2	DP3	Total
Year 7	~~3924~~	~~2628~~	883	**7435**
Year 8	3034	1958	~~2508~~	**7500**
Year 9	2710	1798	1646	**6154**
Year 10	2127	1531	1514	**5172**
Year 11	1835	768	~~1343~~	**3946**
Year 12			143	**143**
Total	**13630**	**8683**	**8037**	**30350**

Cohort @Start	Start Year	End Year	Start New Year	Total
Year 7	3924	2628	2508	**9060**
Year 8	3034	1958	1646	**6638**
Year 9	2710	1798	1514	**6022**
Year 10	2127	1531	1342	**5000**
Year 11	1835	768	144	**2747**
Total	**13630**	**8683**	**7154**	**29467**

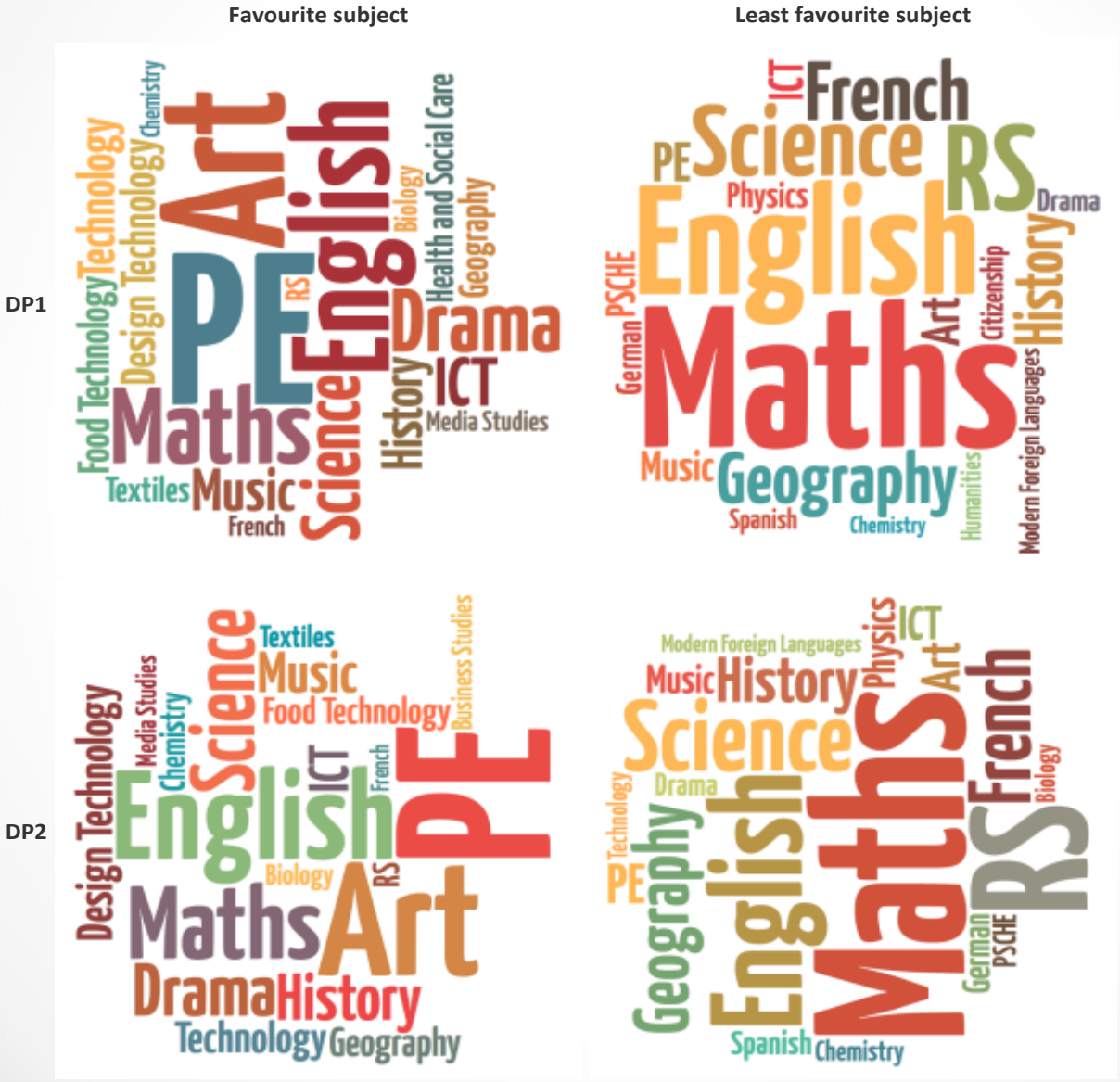
Challenge 1: Attrition and dealing with missing data

More complications….

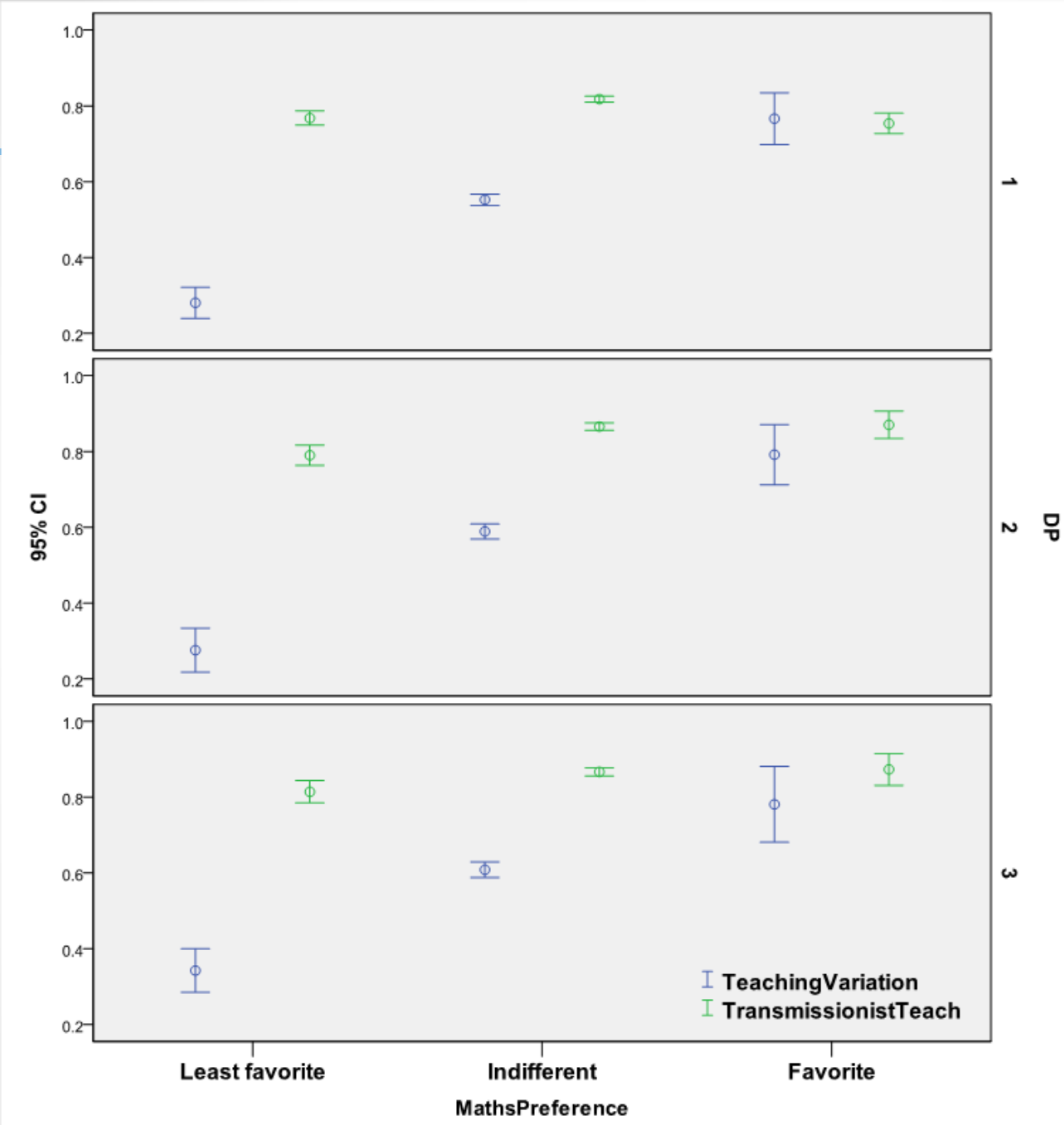
- School level patterns of completion/participation
- School level attrition

School name	dp1	dp2	dp3
School 1	837	738	471
School 2	132		49
School 3	382		386
School 4	224	207	108
School 5	12		
School 6	682		
School 7	170	40	
School 8	686	584	170
School 9	186	187	
School 10	405	267	15
School 11	134	103	
School 12	182	138	125
School 13	69	58	71
School 14	1103	667	602
School 15	128	53	
School 16	748	730	553
School 17	584	585	493
School 18	179	182	202
School 19	261	216	248
School 20	764	675	512
School 21	145	137	
School 22	635	569	596
School 23	45		
School 24	353	28	191
School 25	128	109	98
School 26	136	111	202
School 27	715	615	727
School 28	154	139	
School 29	548		492
School 30	105	106	
School 31	59		
School 32	341		
School 33	150	143	153
School 34	678	627	517
School 35	123	111	85
School 36	28	29	23
School 37	283		
School 38	435	436	441
School 39	167	165	141
School 40	420		

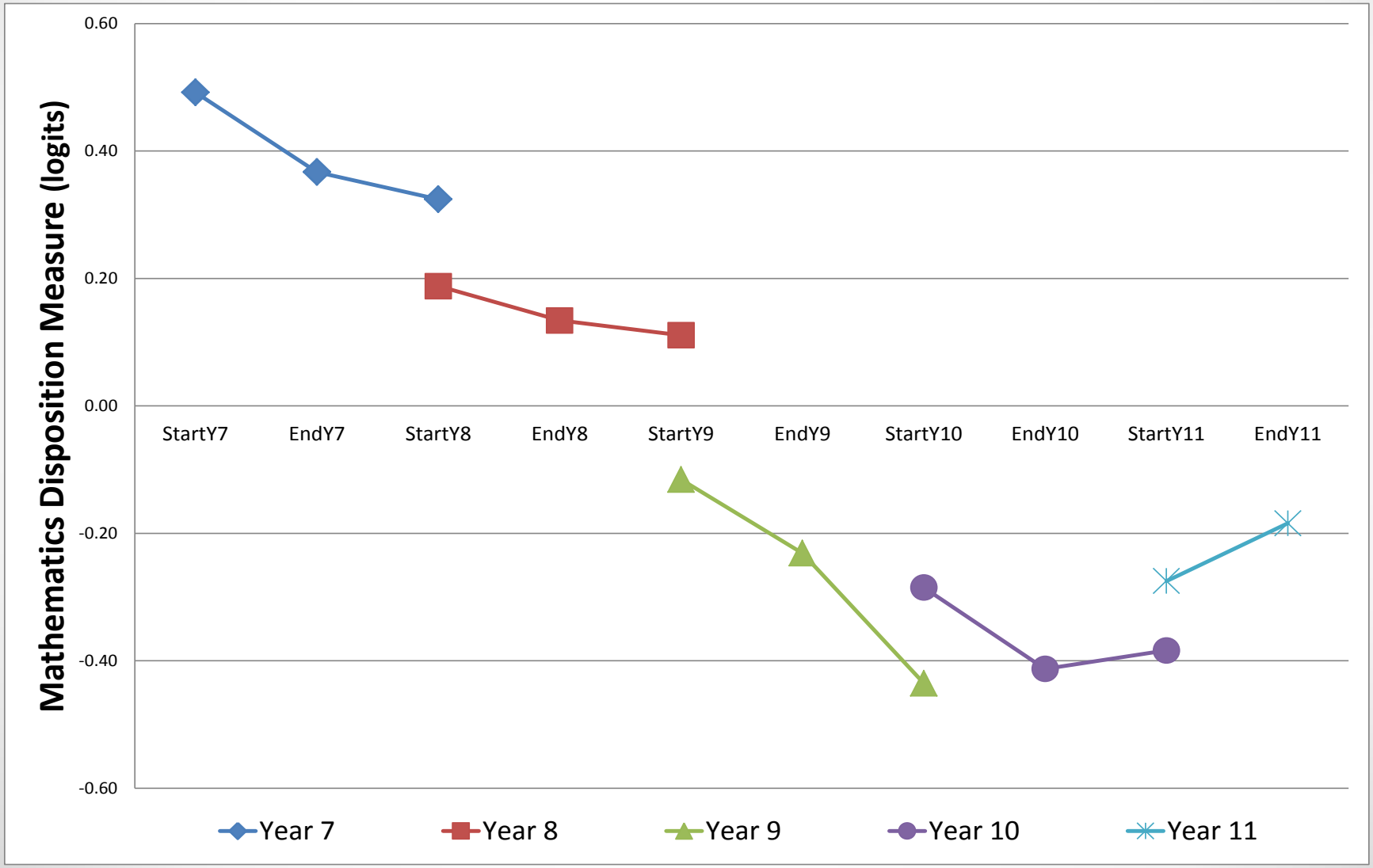
Some descriptive results: favorite topics



Association between Maths preference and teaching style

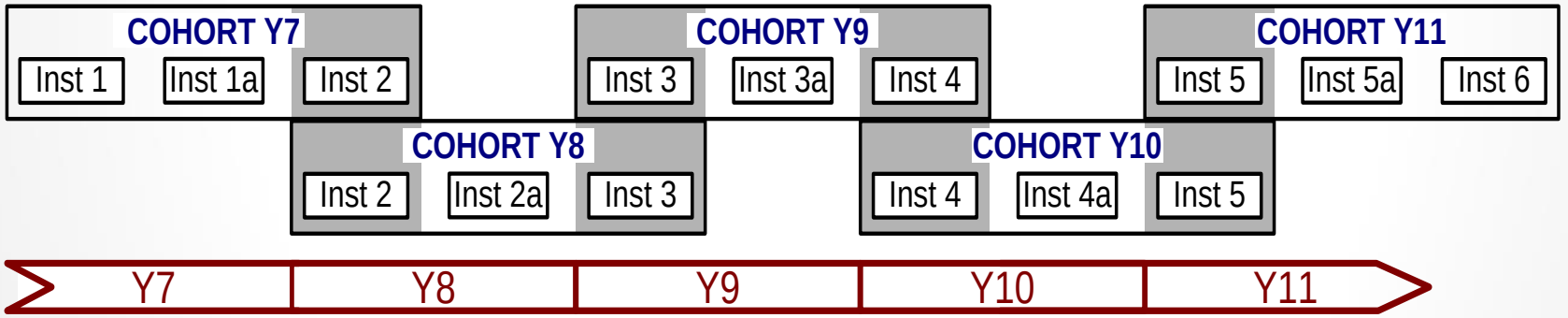


Mean Maths disposition, by cohort



Analytical challenge

RQ3: How can cross-sectional and longitudinal models be combined in the context of hierarchical data structures and missing data?



Accelerated longitudinal design

Relevant methodological Literature suggestions

- Longitudinal analysis under the multilevel framework
  - Growth curve (latent trajectory) models
  - Dynamic (autoregressive) models
  - Fixed, random, mixed effects models …
- For accelerated designs
  - "the growth curve is estimated on a combination of longitudinal and cross-sectional information" (Hox, 2010, p. 110)
  - Suggested Procedure: each cohort analysed separately and then combined (formulation and testing of 'linkage model')
- Further Complications
  - Most examples/applications for 2-level "occasion within subject"
  - Here multilevel in schools (and classes) as well
  - No class information at DP3 (new academic year)

Preliminary Modelling (so far)

- 2-level longitudinal models (level 1: DP, level 2: student)
  - xtreg (stata)
  - xtmixed (stata)
  - runmlwin (mlwin within stata)
  - Both by cohort separately and combined
- 3-level growth curve models (DP, students in schools)
  - Separate cohorts
  - combined
- New variable for linkage: age

age	year_cohort 1	2	3	4	5	Total
11	3,924	0	0	0	0	3,924
11.5	2,628	0	0	0	0	2,628
12	2,508	3,034	0	0	0	5,542
12.5	0	1,958	0	0	0	1,958
13	0	1,646	2,710	0	0	4,356
13.5	0	0	1,798	0	0	1,798
14	0	0	1,514	2,127	0	3,641
14.5	0	0	0	1,531	0	1,531
15	0	0	0	1,342	1,835	3,177
15.5	0	0	0	0	768	768
16	0	0	0	0	144	144
Total	9,060	6,638	6,022	5,000	2,747	29,467

- Treat cohorts as dummy variables
- Estimate fixed effects in the form of interactions with cohort (Plewis, 2009)

MLwiN 2.30 multilevel model Number of obs = 28130
Normal response model
Estimation algorithm: IGLS

Level Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
school_id	40	12	703.3	2145
unique_use~r	16614	1	1.7	3

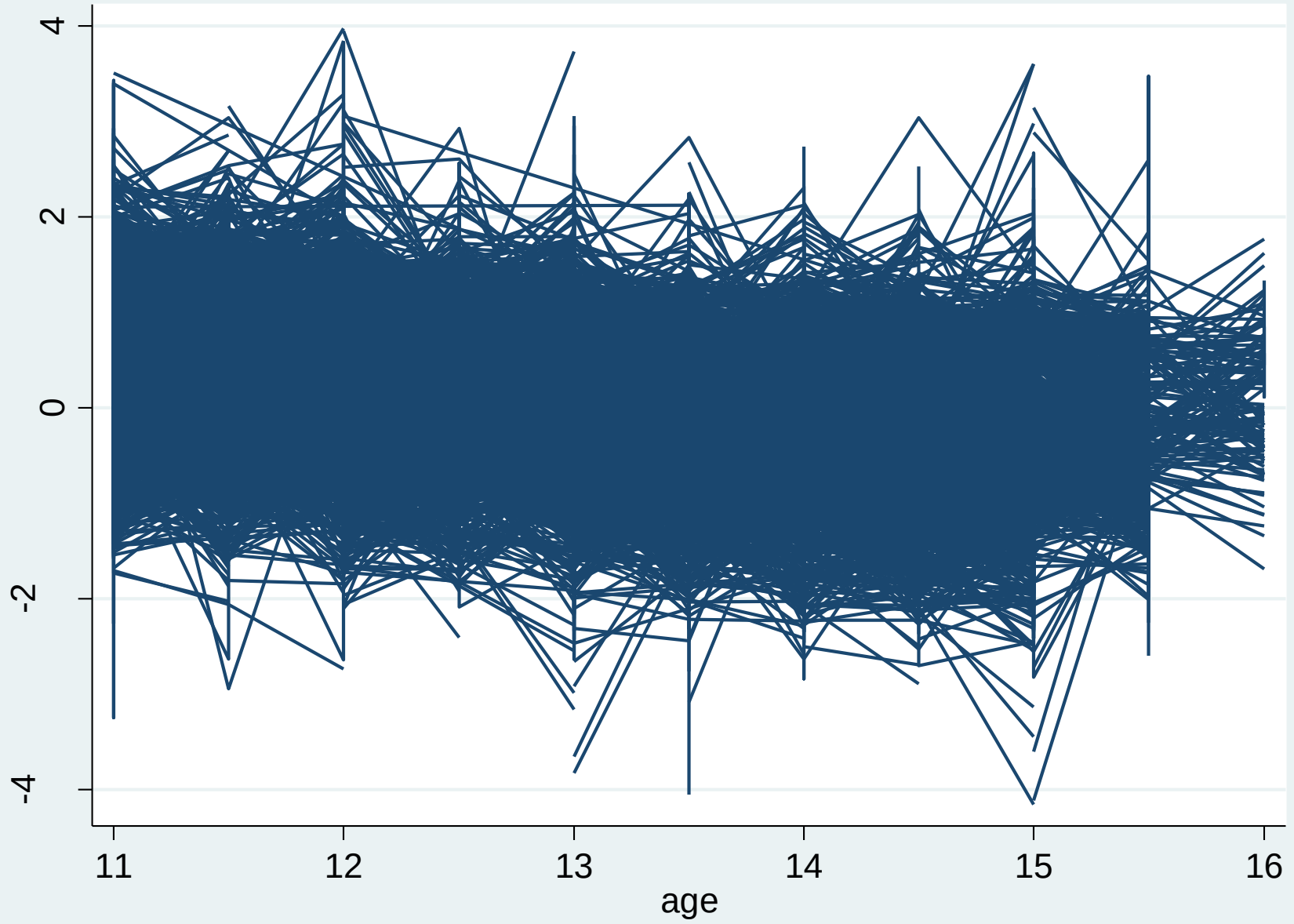
Run time (seconds) = 4.53
Number of iterations = 3
Log likelihood = -44995.602
Deviance = 89991.203

mathsdisposition	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
cons	.2654768	.1892702	1.40	0.161	-.1054859	.6364396
age	-.1088022	.0160247	-6.79	0.000	-.1402099	-.0773944
_Iyear_coh~2	-.021217	.0310303	-0.68	0.494	-.0820352	.0396012
_Iyear_coh~3	-.1110374	.0415801	-2.67	0.008	-.1925328	-.029542
_Iyear_coh~4	-.0480477	.0557697	-0.86	0.389	-.1573544	.061259
_Iyear_coh~5	-.00816	.0693778	-0.12	0.906	-.144138	.1278179
_Igender_2	-.233826	.02014	-11.61	0.000	-.2732996	-.1943523
Iability~2	.7430826	.0352745	21.07	0.000	.6739458	.8122194
Iability~3	1.426746	.0353747	40.33	0.000	1.357412	1.496079
Iability~4	2.31953	.0385862	60.11	0.000	2.243902	2.395158
parentalsu~t	.1588398	.0076943	20.64	0.000	.1437592	.1739203
TeachingVa~n	.3169403	.0089465	35.43	0.000	.2994055	.334475
TeachingTr~e	-.1162968	.0189731	-6.13	0.000	-.1534833	-.0791102

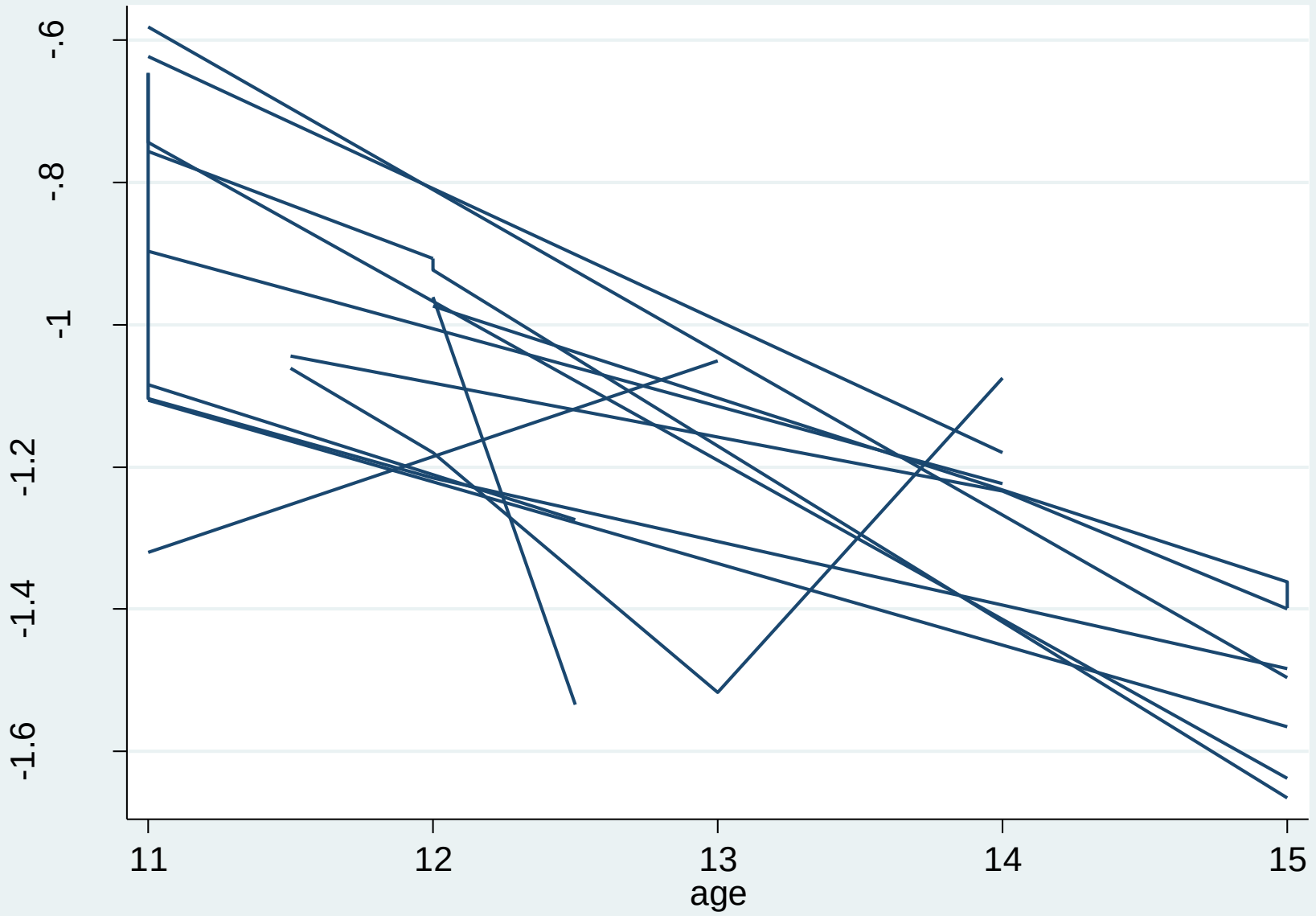
Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
Level 3: school_id var(cons)	.0501309	.0126577	.0253222	.0749396
Level 2: unique_user_number var(cons)	.7319179	.0156322	.7012793	.7625565
Level 1: dp var(cons)	.8630708	.0110958	.8413234	.8848182

- Level 3 Variance (school): 0.0305
- Level 2 Variance (student): 0.4449
- Level 1 Variance (time): 0.5246

‘Growth curves’ of students



‘Growth curves’ of schools



Concluding points / Further work

- Evidence of declining dispositions
- Effect of teaching style on decline
- Possible to link and model progression from Year 7 to 11

- Improve and test the models (age*cohort interactions, etc)
- Additional complexity (levels and variables):
  - Class level (for first year: DP1 and DP2) and teacher background and teaching style
  - Cross-level interactions ?
  - More student background variables
  - School level variables
- Non –linear growth?
- Missing data
- Comparability with national picture (with NPD)

Thank you!

Questions?

Suggestions welcome!