Language requirements for permanent residence & citizenship

The impact on low-educated, low-literate migrants

Bart Deygers, FWO & KU Leuven – Belgium

244 000 000 migrants internationally

International Organization for Migration. (2017). *World migration report 2018*. Geneva: IOM.

75 000 000 in Europe

International Organization for Migration. (2017). *World migration report 2018*. Geneva: IOM.

25 000 000 limited access to formal education

International Organization for Migration. (2017). *World migration report 2018*. Geneva: IOM.

A neglected research population

Most research is premised on WEIRD participants

(Henrich et al., 2010; Ortega 2005, 2019; Tarone & Bigelow, 2012)

A neglected research population

Most research is premised on WEIRD participants

(Henrich et al., 2010; Ortega 2005, 2019; Tarone & Bigelow, 2012)

7% of the global adult population has a university degree

(Barro & Lee, 2013)

A neglected research population

Most research is premised on WEIRD participants

(Henrich et al., 2010; Ortega 2005, 2019; Tarone & Bigelow, 2012)

14% of the global adult population is illiterate

(UNESCO, 2017)

A neglected research population

We do not know how to *teach* LESLLA learners

(OECD, 2018)

A neglected research population

We don't really know how to *teach* LESLLA learners

(OECD, 2018)

But we do know

… that the amount of hours of instruction provided is mostly insufficient

(Kurvers, 2015; Malessa, 2018; UNESCO, 2018; Schellekens, 2011)

A neglected research population

We don't really know how to *teach* LESLLA learners

(OECD, 2018)

But we do know

… that the amount of hours of instruction provided is mostly insufficient

(Kurvers, 2015; Malessa, 2018; UNESCO, 2018; Schellekens, 2011)

… that L2 courses are not as efficient as typically projected

(De Niel et al, 2016; Schuurmans, 2008)

A neglected research population

We don't really know how to *test* LESLLA learners

A neglected research population

We don't really know how to *test* LESLLA learners

(Allemano, 2013; Carlsen, 2017)

But we do know that alphabetic literacy impacts

phonemic awareness

(Dehaene et al., 2010)

working memory capacity

(Demoulin & Kolinsky, 2016; Huettig & Mishra, 2014)

processing speed

(Bengtsson et al., 2005)

A neglected research population

We do not know how to *test* LESLLA learners

(Allemano, 2013; Carlsen, 2017)

But we do know that that schooling impacts basic test-taking strategies & problem-solving abilities

(Allemano, 2013; Oller, Kim, & Choe, 2000; Ostrosky-Solis et al. , 1998)

We do not know how to *test* LESLLA learners

But we do so anyway

Three main questions

Study 1: How often does it happen?

Study 2: How do LESLLA learners perform?

Study 3: Does teaching help?

Fil rouge: fairness & justice

Study 1: How often does it happen

Is the policy justifiable?

Study 2: How do LESLLA learners perform?

Is the test fair?

Study 3: Does teaching help?

Does teaching even the odds?

Fil rouge: fairness & justice

Fairness: Test-internal, test quality (rater severity, bias…)

Justice: Text-external, policy-related

(McNamara, Knoch & Fan, 2019; McNamara & Ryan, 2011)

Study 1: How often does it happen?

Study 2: How do LESLLA learners perform?

Study 3: Does teaching help?

Often.

> 50% of the countries worldwide

(MIPEX 2015)

Council of Europe

Democracy
Human rights
The rule of law
Non-discrimination
Freedom of expression
…

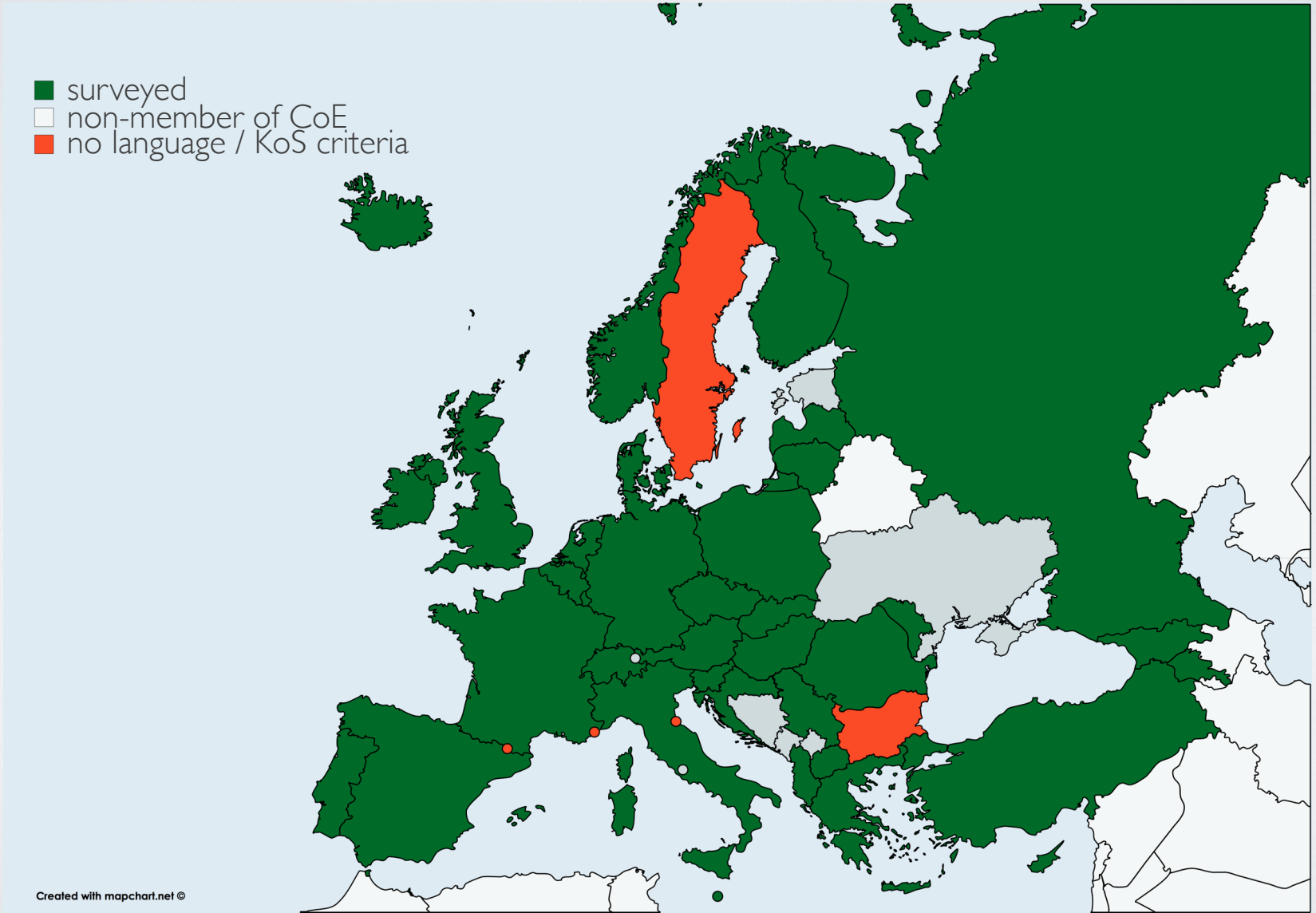
°1949, currently 47 member states

Survey

Language requirements for migration, residence and citizenship

2007: 27 member states (Little, 2008)
2009: 34 member states (Extramiana & Van Avermaet, 2010)
2013: 37 member states (CoE Language Policy Unit, 2014)
2018: 41 member states (in press)

Use of language criteria



Research-based?

"Language experts from [government department] set the levels"

"The language requirements are mainly the result of politically motivated decisions"

Research-based?

19 January 2017 at 10:09

LV

Testen A2: onderzoek -- negatief antwoord van ABB en kabinet

To: Bart Deygers

Hey Bart

Slecht nieuws! Ik heb van te horen gekregen dat zij na overleg met het kabinet beslist hebben dat het te vroeg is voor onderzoek.
Ik wil wel nog horen wanneer het dan wel zou kunnen, als jij dat zou willen.

Ik persoonlijk zou onze afspraak volgende week wel nog willen laten doorgaan. Laat maar weten of het voor jou nog oké is.

Groetjes
Lieselot

"Bad news! I've been informed that they have decided that it is too early for research."

Research-based?

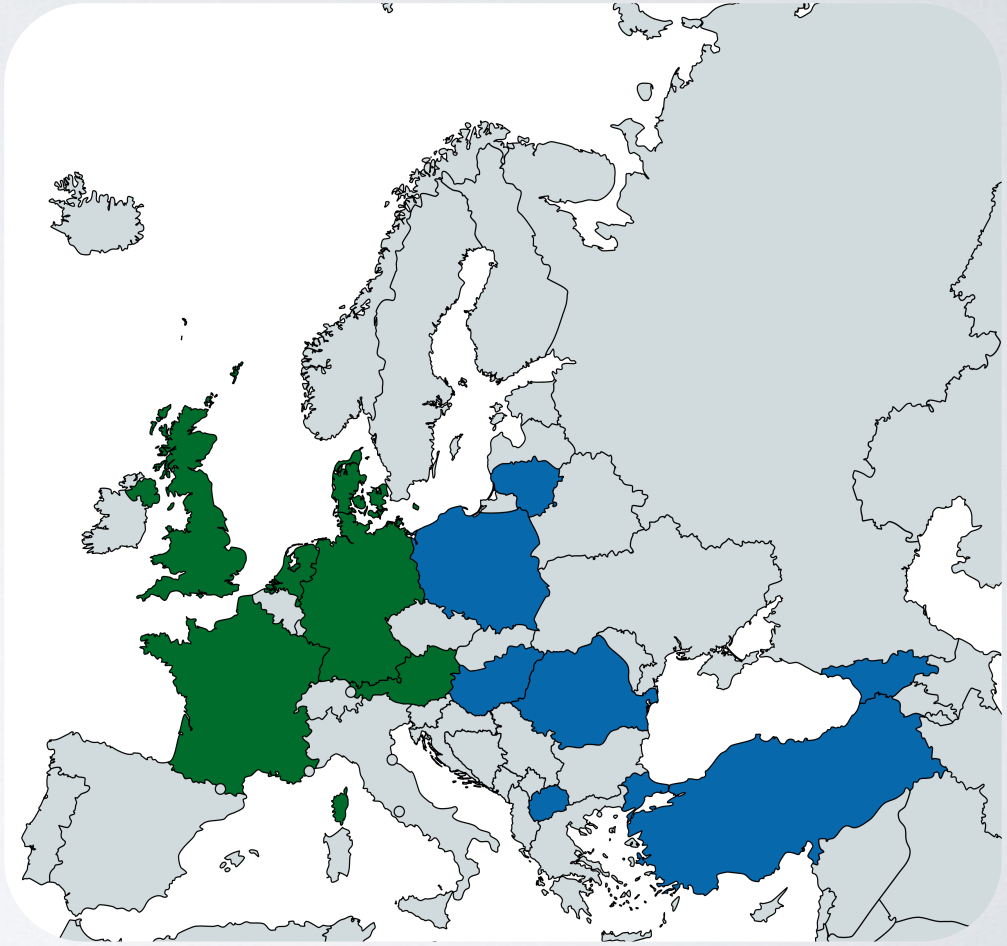
… "the level of proficiency required is not determined by a careful study of the level needed for these purposes, but is used as a lever to control numbers of new permanent residents"

McNamara, Knoch & Fan, 2019, p. 20

(See also: McNamara, Khan, & Frost, 2014; Frost, 2018; Deygers et al., 2018)

Requirements

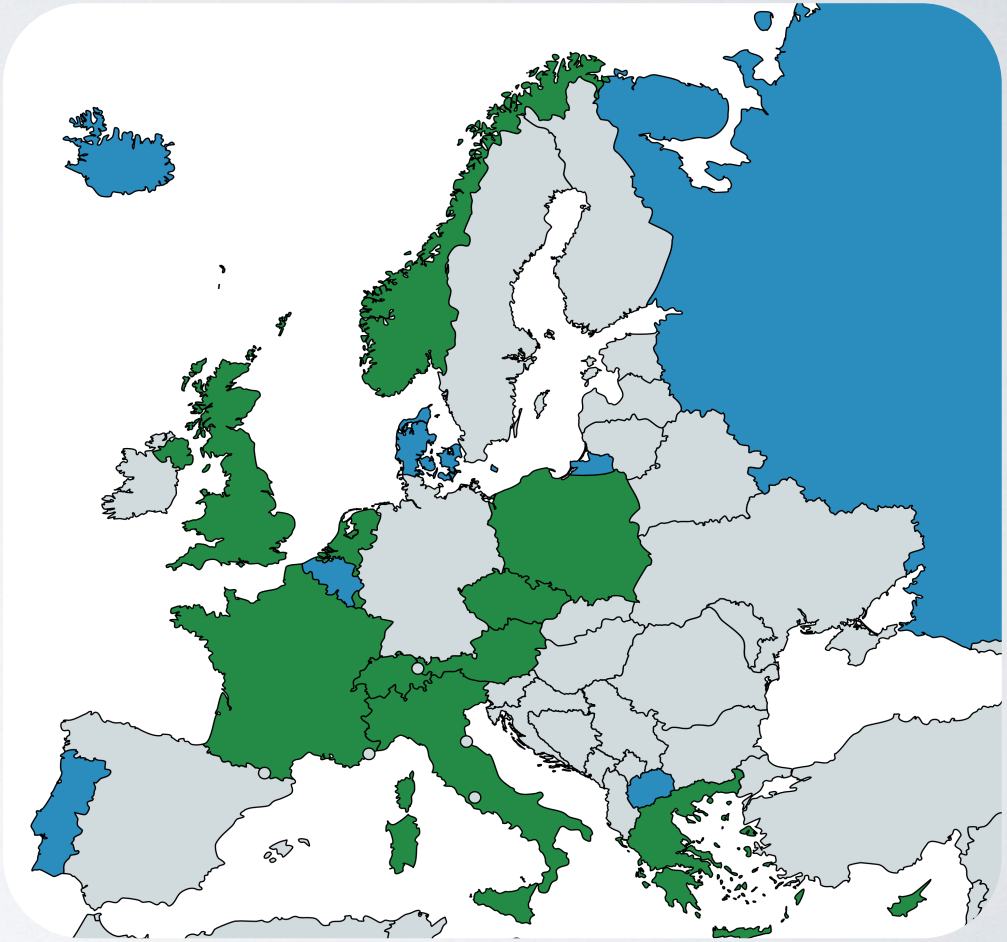
Pre-entry: 13/41
A1



Requirements

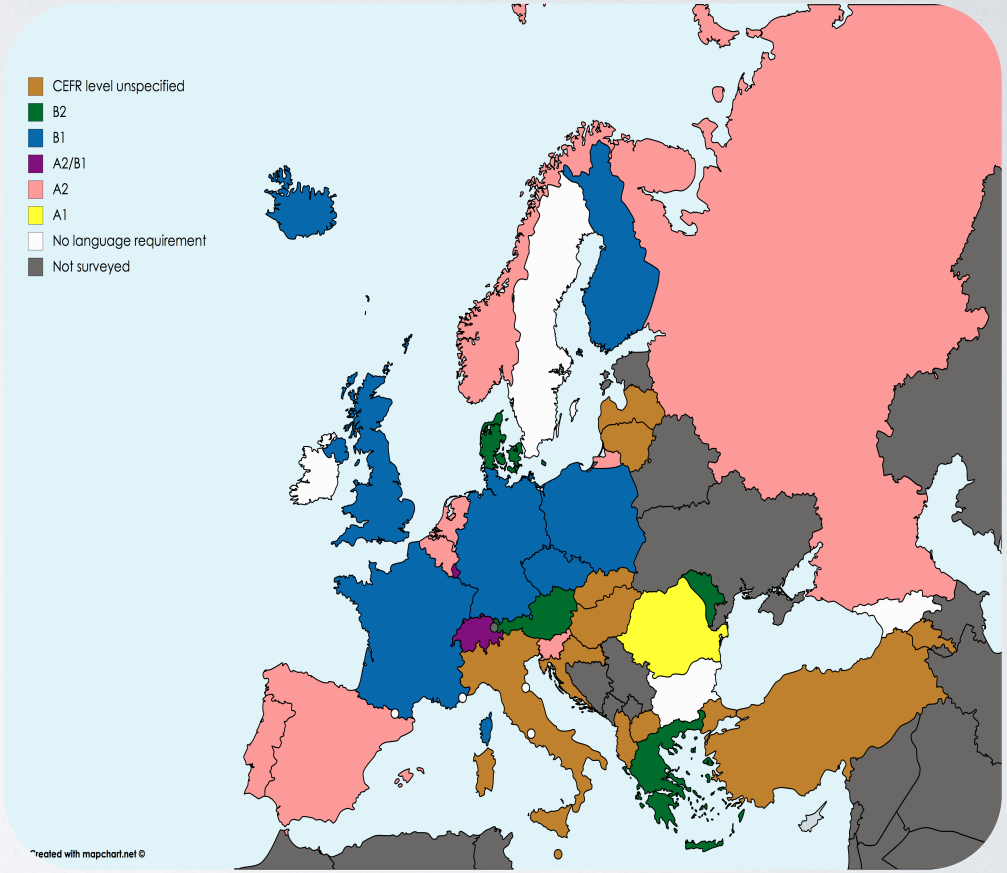
Pre-entry: 13/41
A1

Permanent residence: 23/41
A2



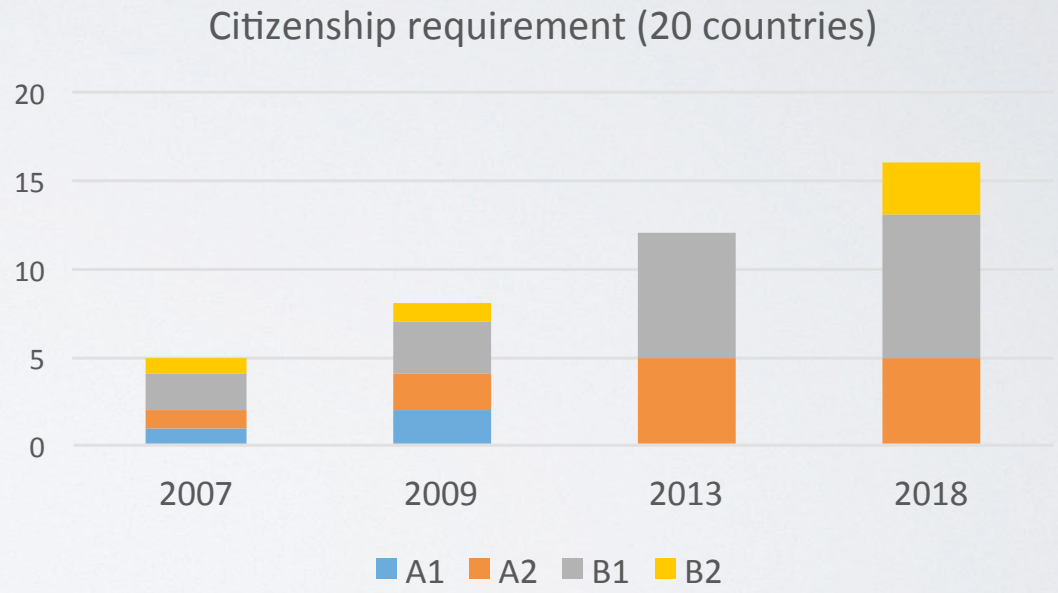
Requirements

Pre-entry:	13/41 A1
Permanent residence:	23/41 A2
Citizenship	32/41 B1



Quick takeaways

Proportional increase in requirements
Gradual increase in level (A1 – A2 – B1)



Quick takeaways

Proportional increase in requirements
Gradual increase in level
Increased use of language tests

BUT External quality control: 7 countries

Quick takeaways

Proportional increase in requirements
Gradual increase in level
Increased use of language tests

Sharp and significant increase in Knowledge of Society tests (9/37 – 16/41)

And LESLLA?

Hardly any exemptions from language or KoS requirements

And LESLLA?

Hardly any exemptions from language or KoS requirements

20 countries do offer specific language courses

And LESLLA?

Hardly any exemptions from language or KoS requirements

20 countries do offer specific language courses

But typically just 250 hours of instruction

Can we justify this?

Rawls, 1999

"Peoples have a duty to assist other peoples living under unfavourable conditions that prevent their having a just or decent political and social regime"

(Rawls, J. (1999). *The Law of Peoples*, p. 10)

Can we justify this?

Rawls, 1999

Shohamy, 2007

> Language proficiency can never be a proxy for good citizenship (whatever that may be)

Can we justify this?

Rawls, 1999
Shohamy, 2007
Valentini, 2011

a system is coercive if it foreseeably and avoidably places non-trivial constraints on some people's freedom, compared to their freedom in the absence of that system

Justification is required (see also Sen)

Can we justify this?

Rawls, 1999
Shohamy, 2007
Valentini, 2011
Council of Europe, 2013

Certain language requirements in migration policies constitute a breach of fundamental human rights (e.g., pre-entrance requirements and family reunification)

Can we justify this?

Rawls, 1999
Shohamy, 2007
Valentini, 2011
Council of Europe, 2013
Deygers, 2017, 2019

A testing policy is *unjust* if it wilfully and avoidably restricts test takers' freedom without an empirically sound or reasonable motivation.

Can we justify this?

Rawls, 1999
Shohamy, 2007
Valentini, 2011
Council of Europe, 2013
Deygers, 2017, 2019
Bruzos, Erdocia & Khan, 2018

Why argue for *better tests* if the practice is unjust?

Can we justify this?

Rawls, 1999
Shohamy, 2007
Valentini, 2011
Council of Europe, 2013
Deygers, 2017, 2019
Bruzos, Erdocia & Khan, 2018
McNamara, Knoch & Fan, 2019

Many instances of language testing for citizenship are unjust.

However

Statism is the de facto world order

(Valentini, 2011)

However

"[the goal is] a world of diversity in which the variety of national cultures finds expression in different sets of citizenship rights, and different schemes of social justice … States should work together to ensure that every community can protect its members' basic rights, but there should be no attempt to impose uniformity."

"David Miller, 'Caney's "International Distributive Justice": A Response', Political Studies, 50 (5) (2002), 974–7, p. 976"

However

Membership of the Australian family is a privilege and should be afforded to those who support our values, respect our laws and want to work hard by integrating and contributing to an even better Australia … we must ensure that our citizenship program is conducted in our national interest

(M. Turnbull, 20 April 2017)

See also: David Miller, 'Caney's "International Distributive Justice": A Response', Political Studies, 50 (5) (2002), 974–7

Three possible responses

The ivory tower strategy

Head-on collision strategy

Collaborative strategy

(Fischer, 2007)
(See also: Deygers & Malone, 2019; LoBianco, ; McNamara, 2009)

On collaboration

Not rejecting a policy maker's premise does not equal compliance

It means getting a seat at the table

and possibly having an impact

(Cf. Fischer, 2003, 2007)

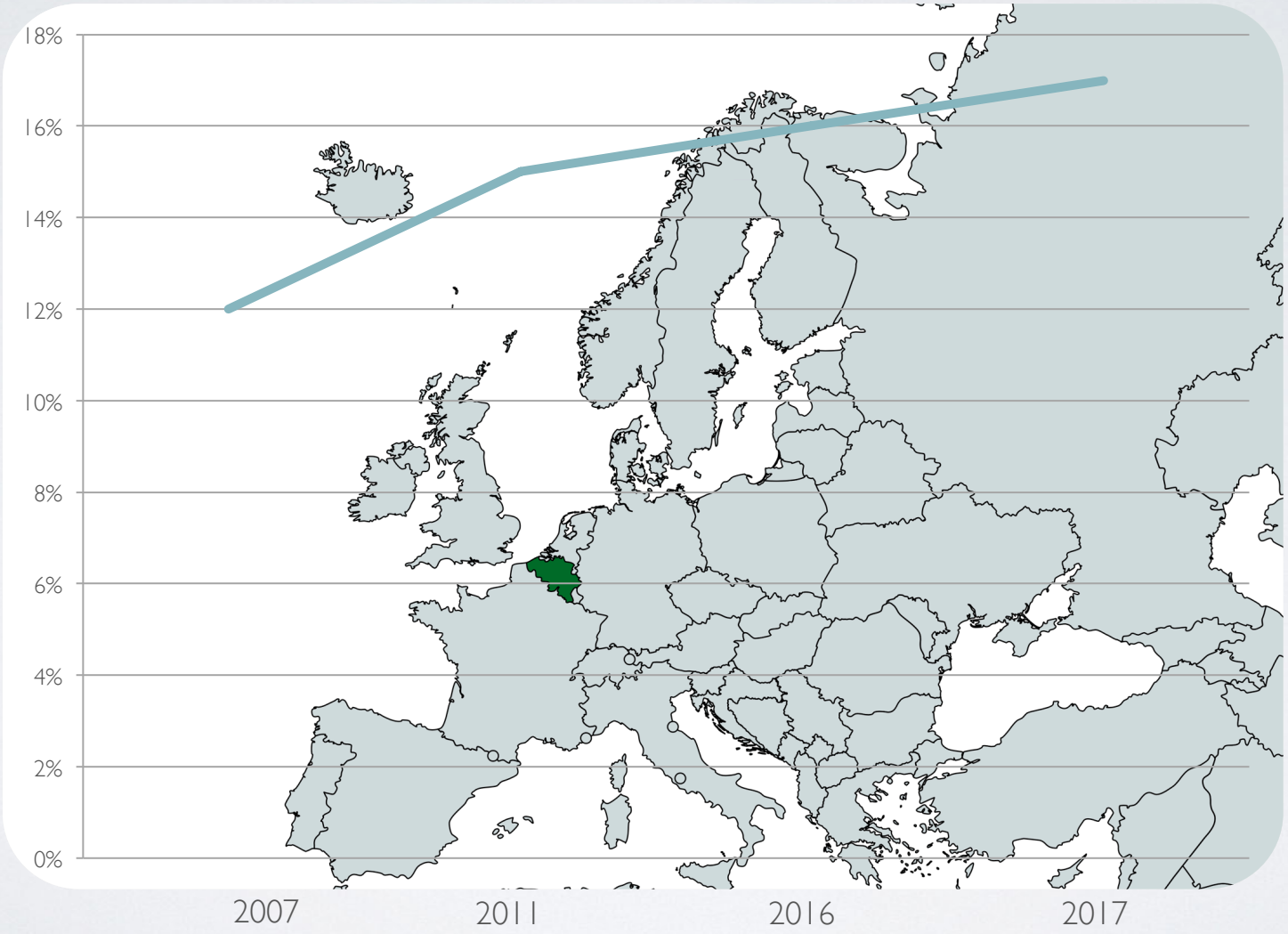
Study 1: LESLLA learners are routinely part of the general population for high-stakes tests

Study 1: LESLLA learners are routinely part of the general population for high-stakes tests

Study 2: How do LESLLA learners perform?

Zoom in on Belgium

Zoom in on Belgium



(OECD, 2018)

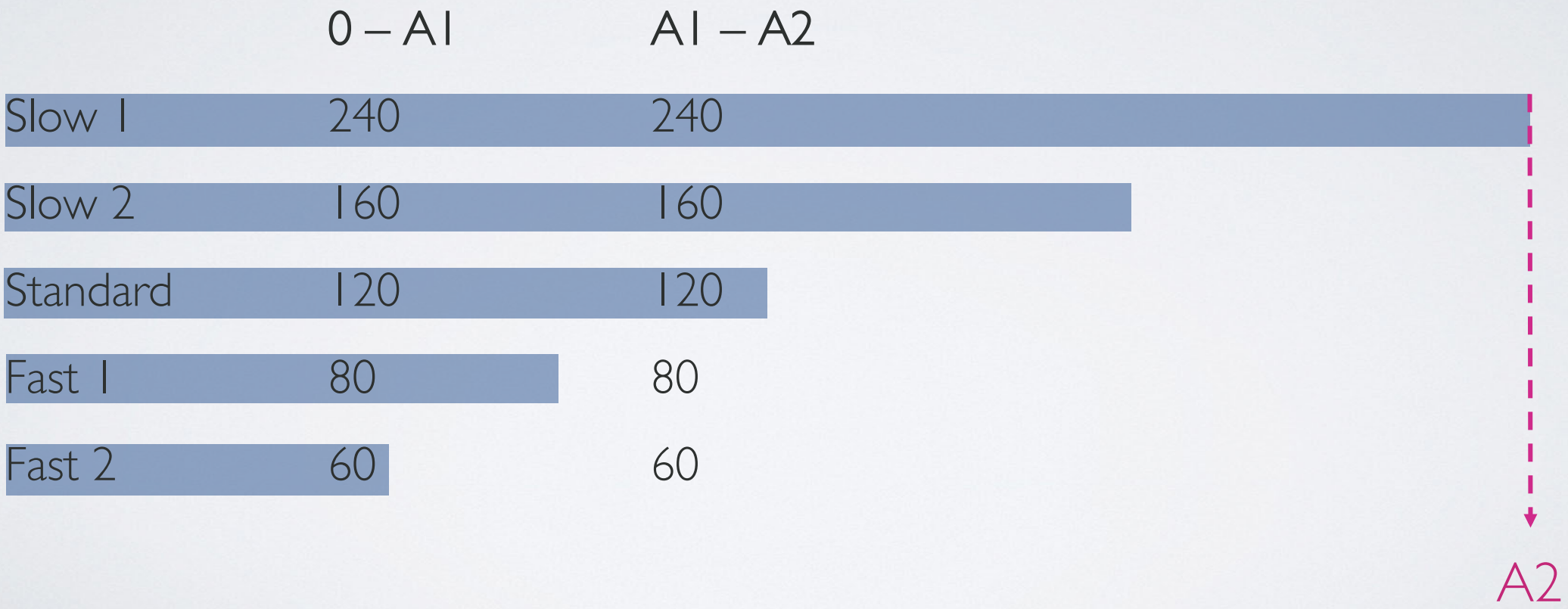
Flanders: language criteria

- 2015 efforts made

Flanders: language criteria

- 2015 efforts made
  2015 - level attained: A2
  2017 A2 language test 

Flanders: language classes



Flanders: language classes

	0 – A1	A1 – A2	
Slow 1	240	240	≥ primary education
Slow 2	160	160	
Standard	120	120	secondary education
Fast 1	80	80	tertiary education
Fast 2	60	60	

Research population

Population *N* = 1058

Age	med 32	mean 34 (se .4, sd 10)
In B	med 2	mean 4 (se .2, sd 4)
52%	female	
25%	employed	

Research population

Population $N = 1058$

A1	54%
A2	46%

Research population

Population *N* = 1058

A1	54%
A2	46%

15%	≤ primary
41%	secondary
28%	tertiary

Alfa	9%
Slow	35%
Standard	37%
Fast	19%

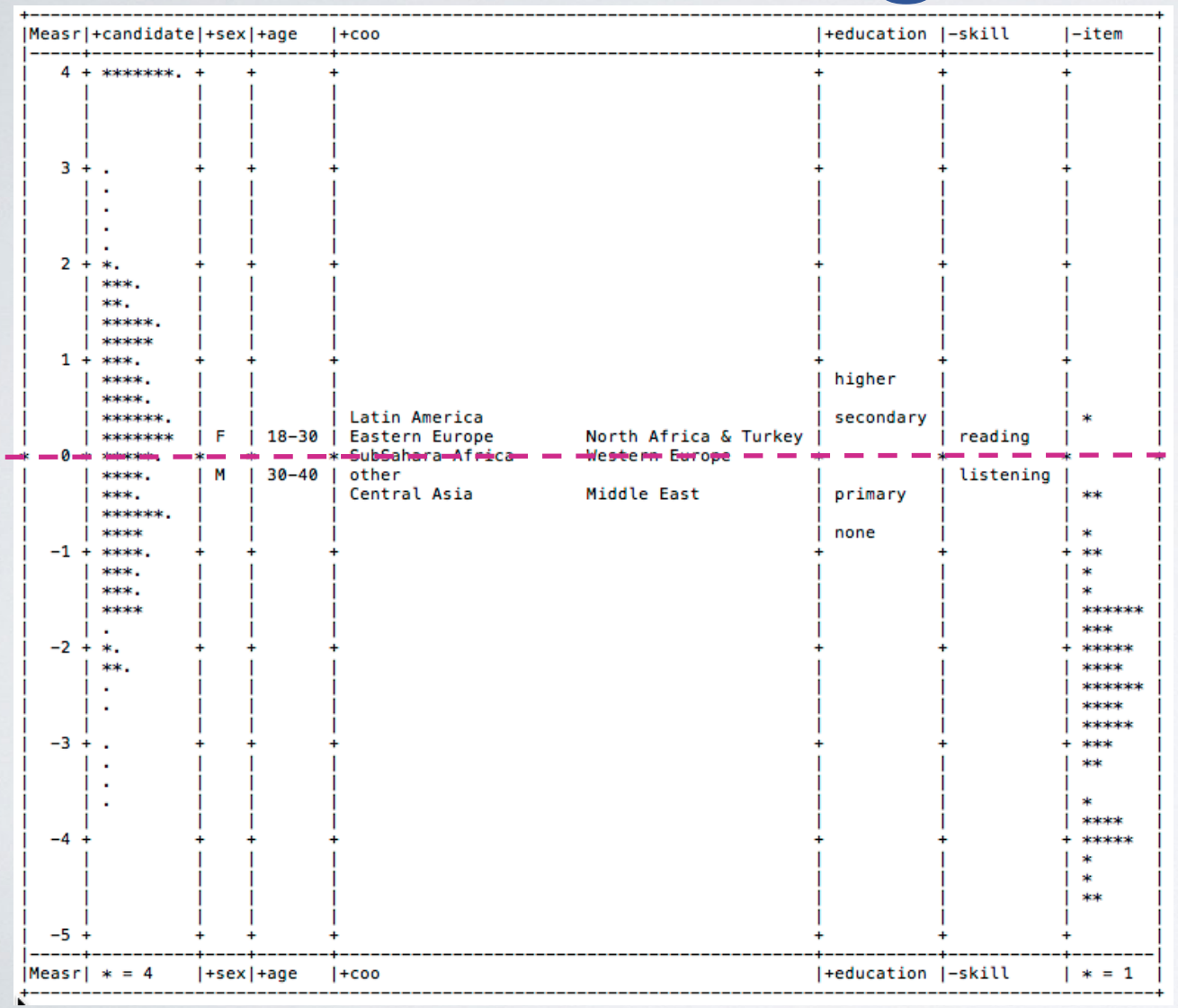
Measurement instruments

Background information survey	(*N* = 1058)
PPVT-III-NL	(*N* = 1058)
Writing test	(*n* = 981 / 385 transcribed and coded)
Speaking test	(*n* = 142 / transcription underway)
Elicited imitation task	(*n* = 113)

Measurement instruments

Background information survey	(*N* = 1058)
PPVT-III-NL	(*N* = 1058)
Writing test	(*n* = 981 / 385 transcribed and coded)
Speaking test	(*n* = 142 / transcription underway)
Elicited imitation task	(*n* = 113)

Listening: scoring profile



Pass probability listening

$P_{(\leq \text{primary})} = .59$

$P_{(\geq \text{secondary})} = .89$

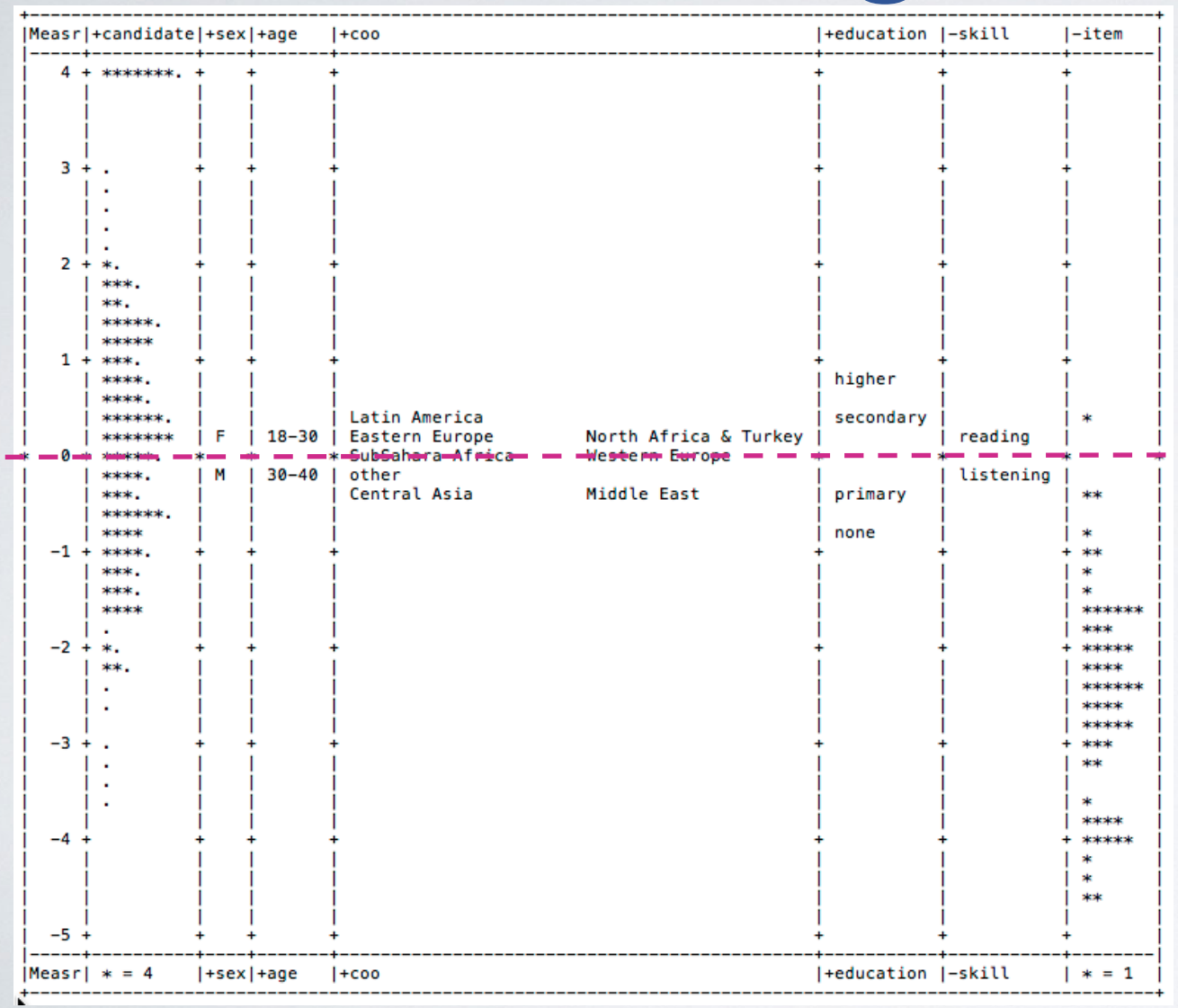
$W = 10040$ / $p < 0.0000$ / d -0.839 / 95% CI -1.084 – -0.594

Listening: score differences

	Median	Lower Secondary	Higher secondary	Higher education
Primary education	25	*W* = 1975.5 *p* <.000 *d* -0.745	*W* = 3417 *p* <.000 *d* -0.940	*W* = 998 *p* <.000 *d* -0.930
Lower secondary	27		W = 6922.5, p = 0.0574 d -0.165	*W* = 2053 *p* = 0.01557 *d* -0.398
Higher secondary	28			*W* = 4995.5 *p* = 0.345 *d* -0.185
Higher education	28			

Plonsky, L., & Oswald, F. L. (2014). How Big Is "Big"? Interpreting Effect Sizes in L2 Research. Language Learning, 64(4), 878–912.

Reading: scoring profile



Pass probability reading

$P_{(\leq \text{ primary})} = .44$

$P_{(\geq \text{ secondary})} = .79$

W = 9251 / *p* < 0.0000 / *d* -0.833 / 95% CI -1.0778 – -0.5878

Reading: score differences

	Median	Lower Secondary	Higher secondary	Higher education
Primary education	21	*W* = 1448 *p* <.000 *d* -1.13	*W* = 3464 *p* <.000 *d* -0.964	*W* = 687 *p* <.000 *d* -1.241
Lower secondary	27		*W* = 9061.5 *p* = 0.091 *d* 0.240	*W* = 1773.5, *p* <.000 *d* -0.376
Higher secondary	27			*W* = 2773 *p* <.000 *d* -0.523
Higher education	29			

Plonsky, L., & Oswald, F. L. (2014). How Big Is "Big"? Interpreting Effect Sizes in L2 Research. Language Learning, 64(4), 878–912.

Listening & reading

Pronounced and significant performance differences depending on educational background ($\chi^2(3) = 370.5, p < .000$)

Listening & reading

Pronounced and significant performance differences depending on educational background ($\chi^2(3) = 370.5, p < .000$)

Educational background substantially impacts score variance

Listening outcome ~ educational background:

B(SE) = 0.203 (0.03), 95% CI 1.226, $p < 0.000$

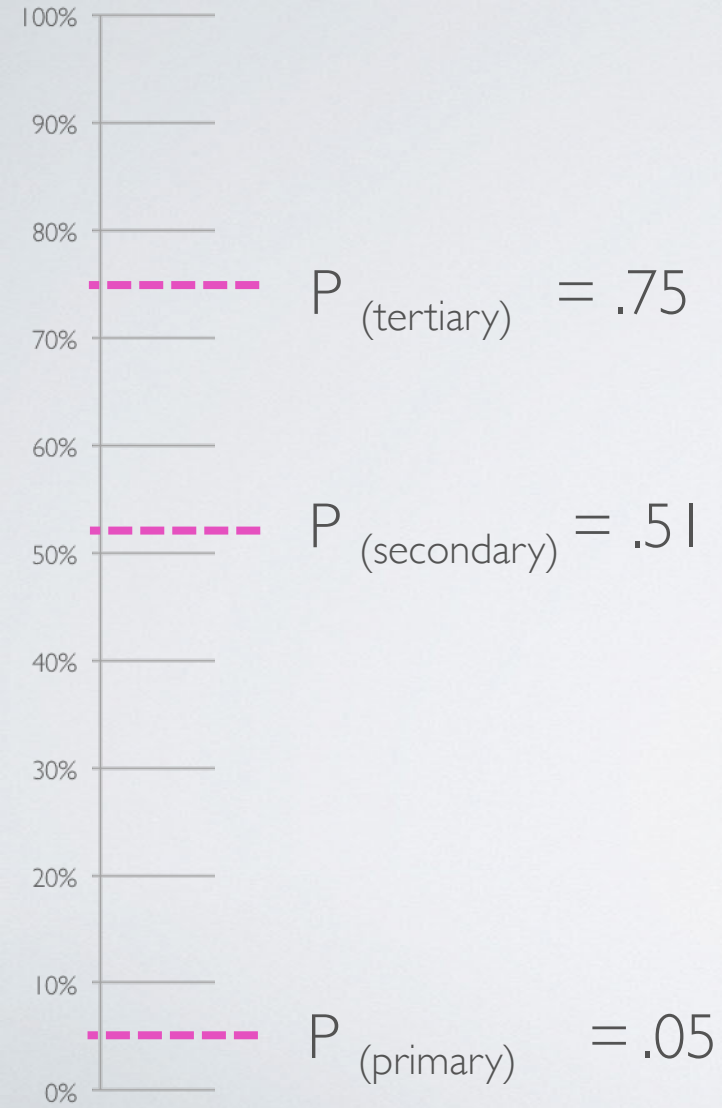
$R^2 = 0.11$ (Nagelkerke)

Reading outcome ~ educational background:

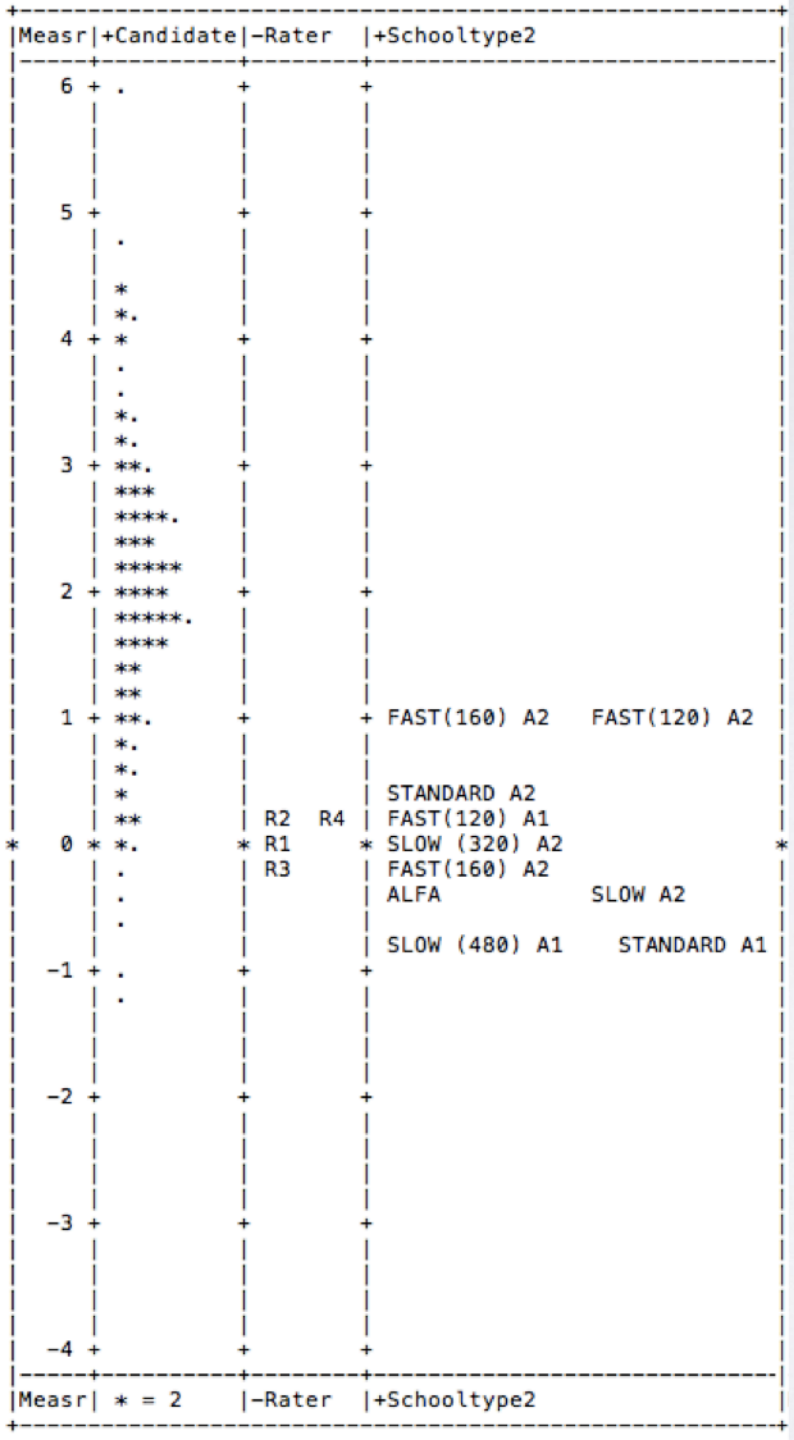
B(SE) = 0.269 (0.034), 95% CI 1.308, $p < 0.0000$

$R^2 = 0.15$ (Nagelkerke)

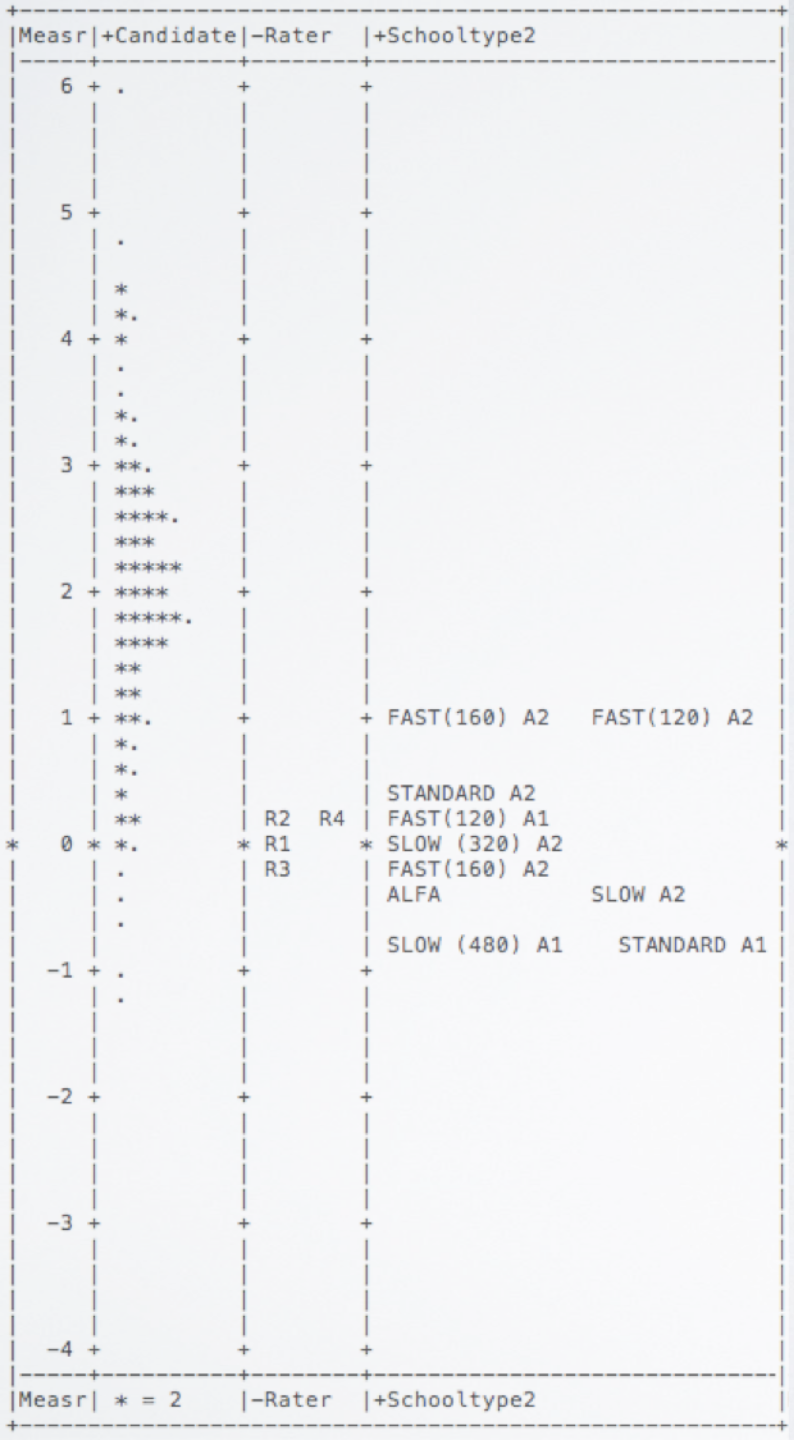
Speaking: pass probability



Primary / Secondary	$W = 816$,	$p = 0.002$; r -0.309
Secondary / tertiary	$W = 1268$,	$p = 0.006$; r -1.194
Primary / tertiary	$W = 508$,	$p = 0.000$; r -1.956



Variance explained by Rasch measures:	83.07%
Variance of residuals:	16.93%
Variance explained by bias/interactions:	
Age	0.64%
Time in Belgium	0.67%
L1	0.91%
Track	2.66%

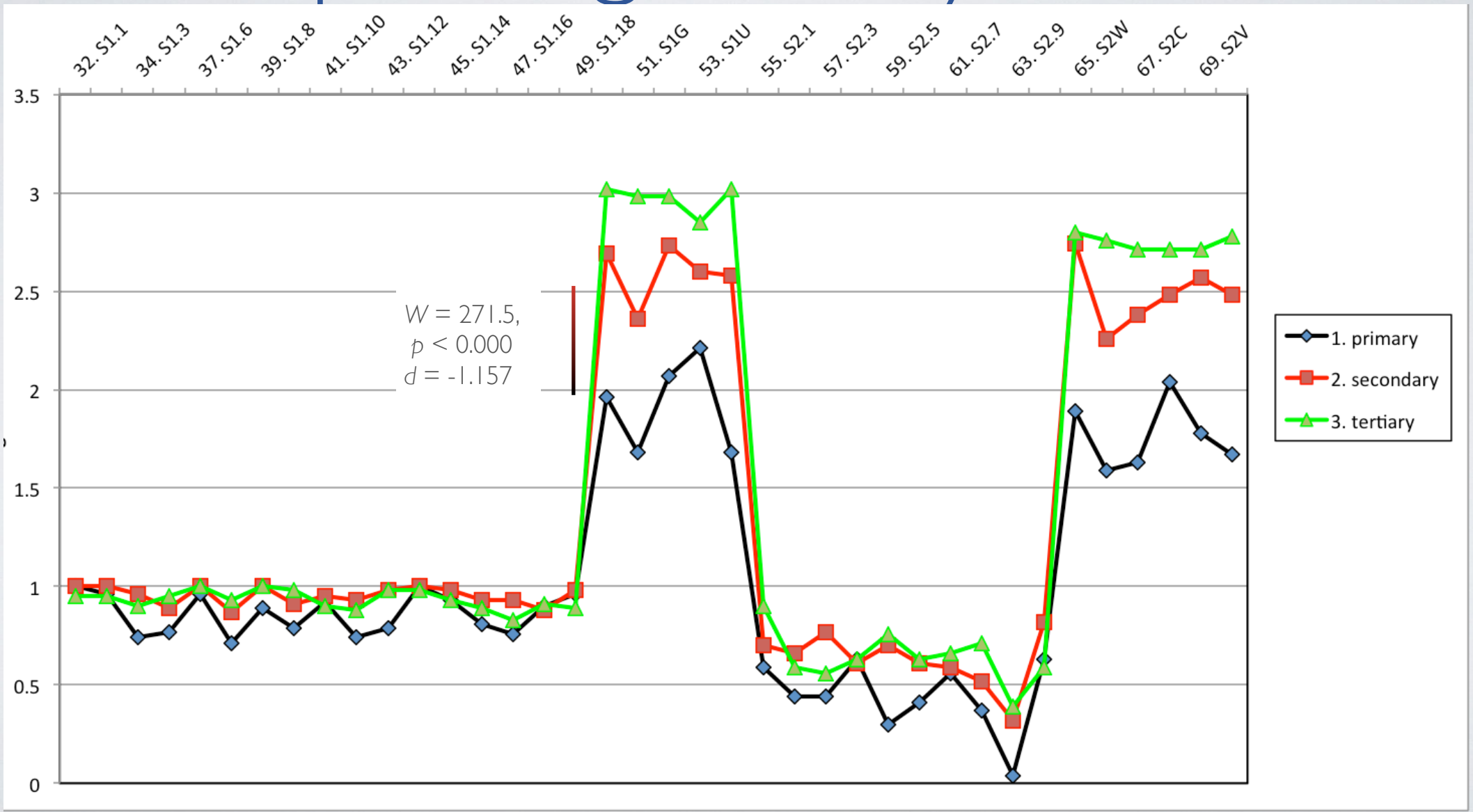


Measure	(se)	Infit	
1.04	0.09	1.14	FAST(120) A2
1.01	0.23	0.84	FAST(160) A2
0.31	0.12	0.84	STANDARD A2
0.24	0.12	1.11	FAST(120) A1
0.04	0.1	0.86	SLOW (320) A2
-0.28	0.14	0.82	FAST(160) A2
-0.44	0.08	0.89	SLOW(480) A2
-0.48	0.11	0.81	ALFA
-0.7	0.11	0.88	STANDARD A1
-0.73	0.07	0.91	SLOW(480) A1

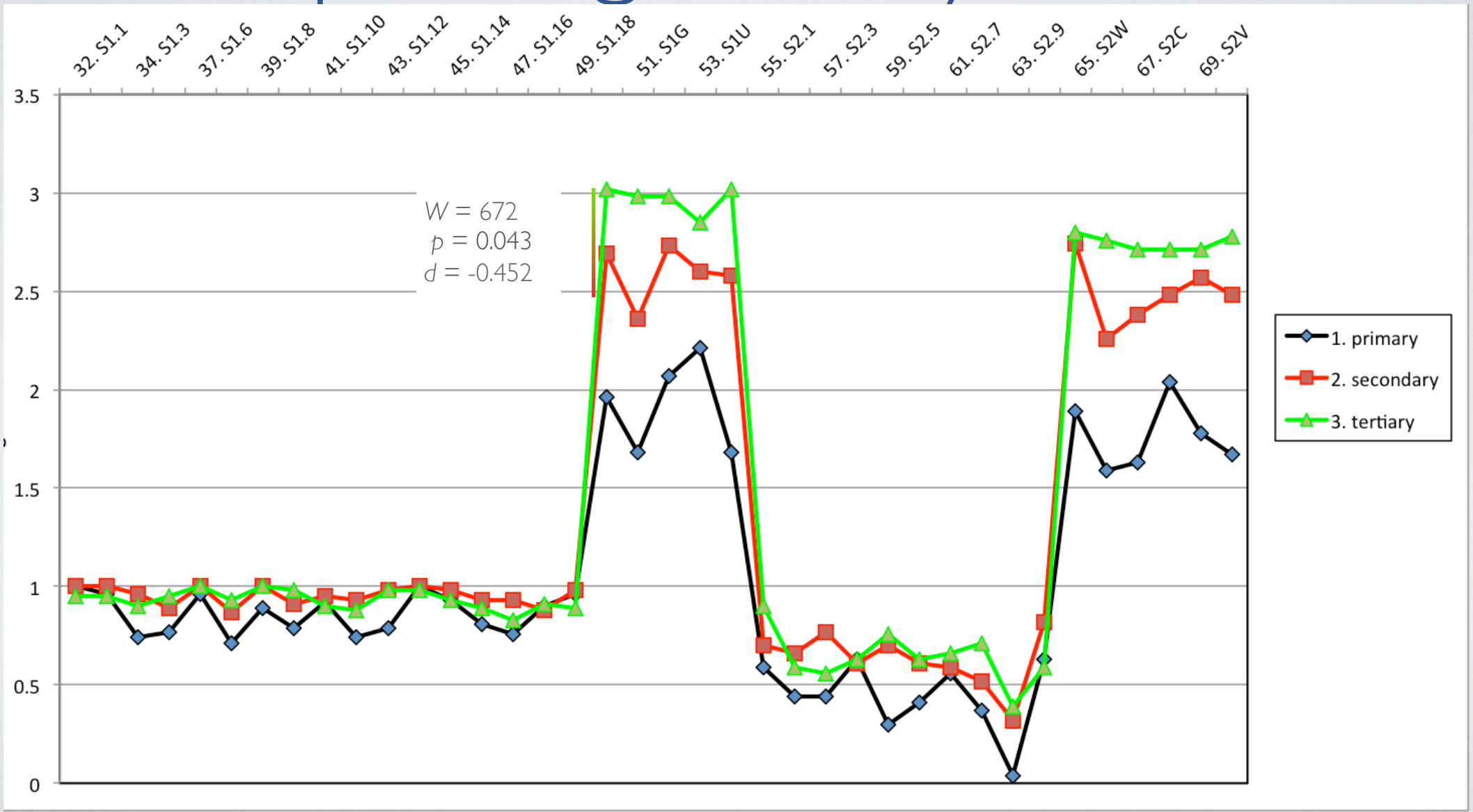
Strata 5.88 Reliability .95

$X^2(9) = 348.5$, $p < .000$

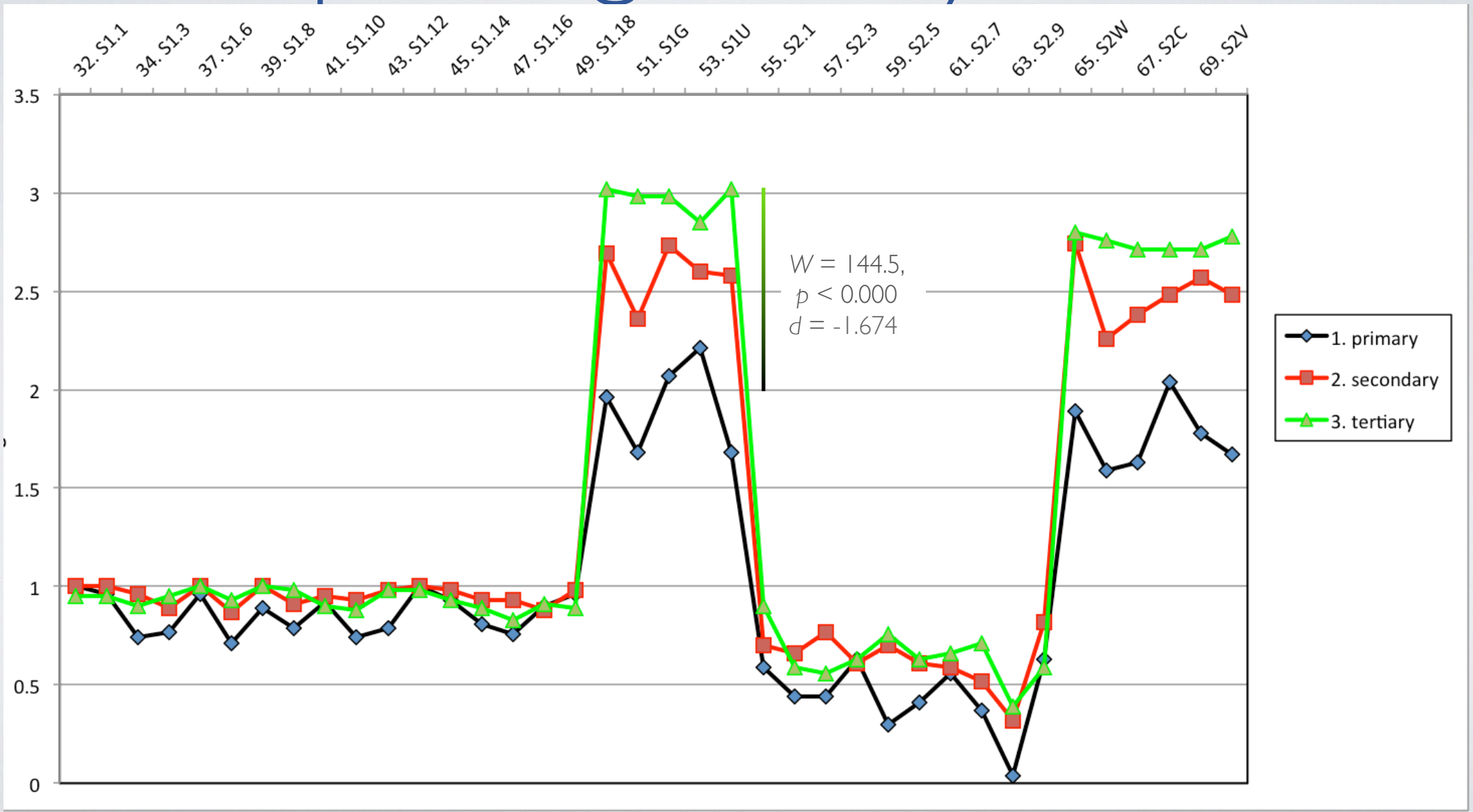
Speaking: bias by item



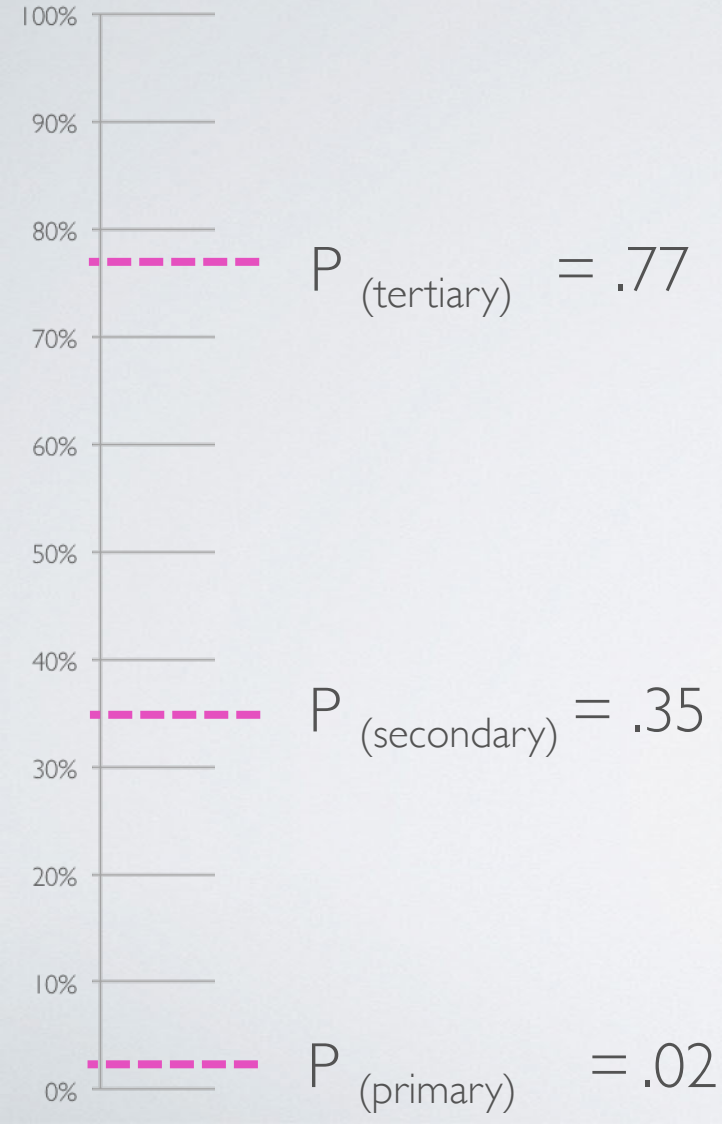
Speaking: bias by item



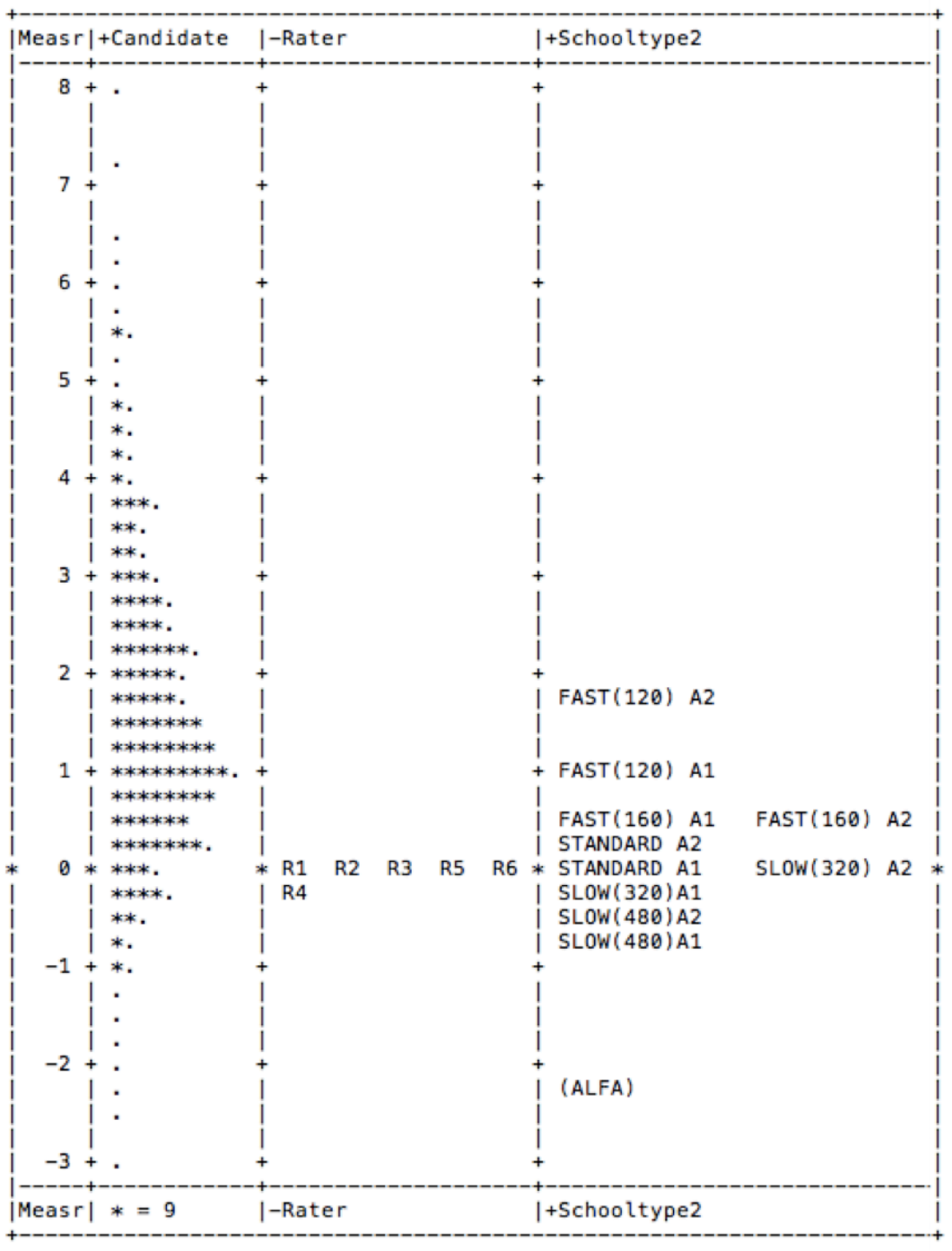
Speaking: bias by item



Writing: pass probability



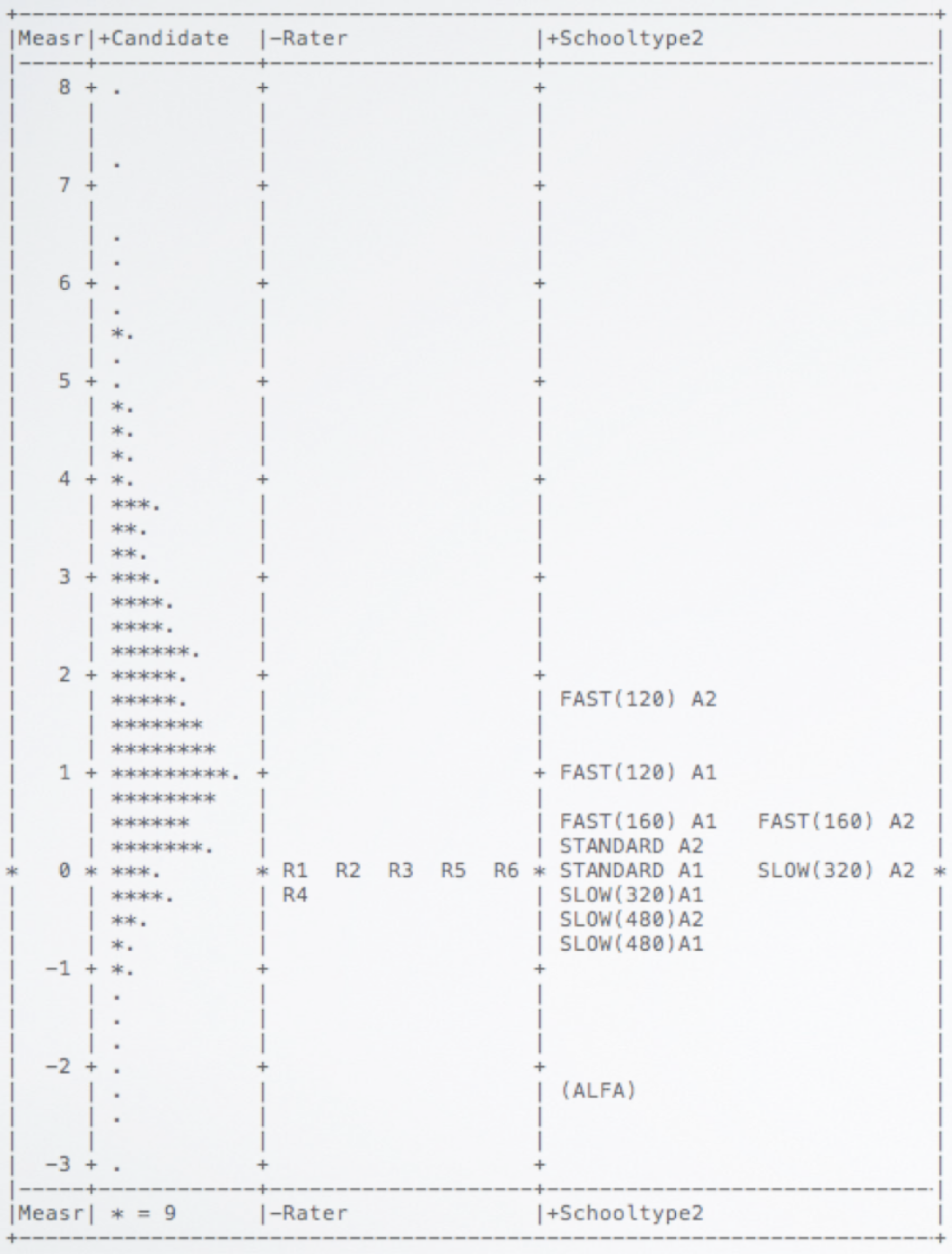
Primary / Secondary	$W = 46256, p < 0.000; r$ -0.367
Secondary / tertiary	$W = 62912, p < 0.000; r$ -0.432
Primary / tertiary	$W = 14893, p < 0.000; r$ -0. 727



Variance explained by Rasch measures:	83.42%
Variance of residuals:	16.58%

Variance explained by bias/interactions:

Age	0.12%
Time in Belgium	0.18%
L1	0.21%
Track	1.20%

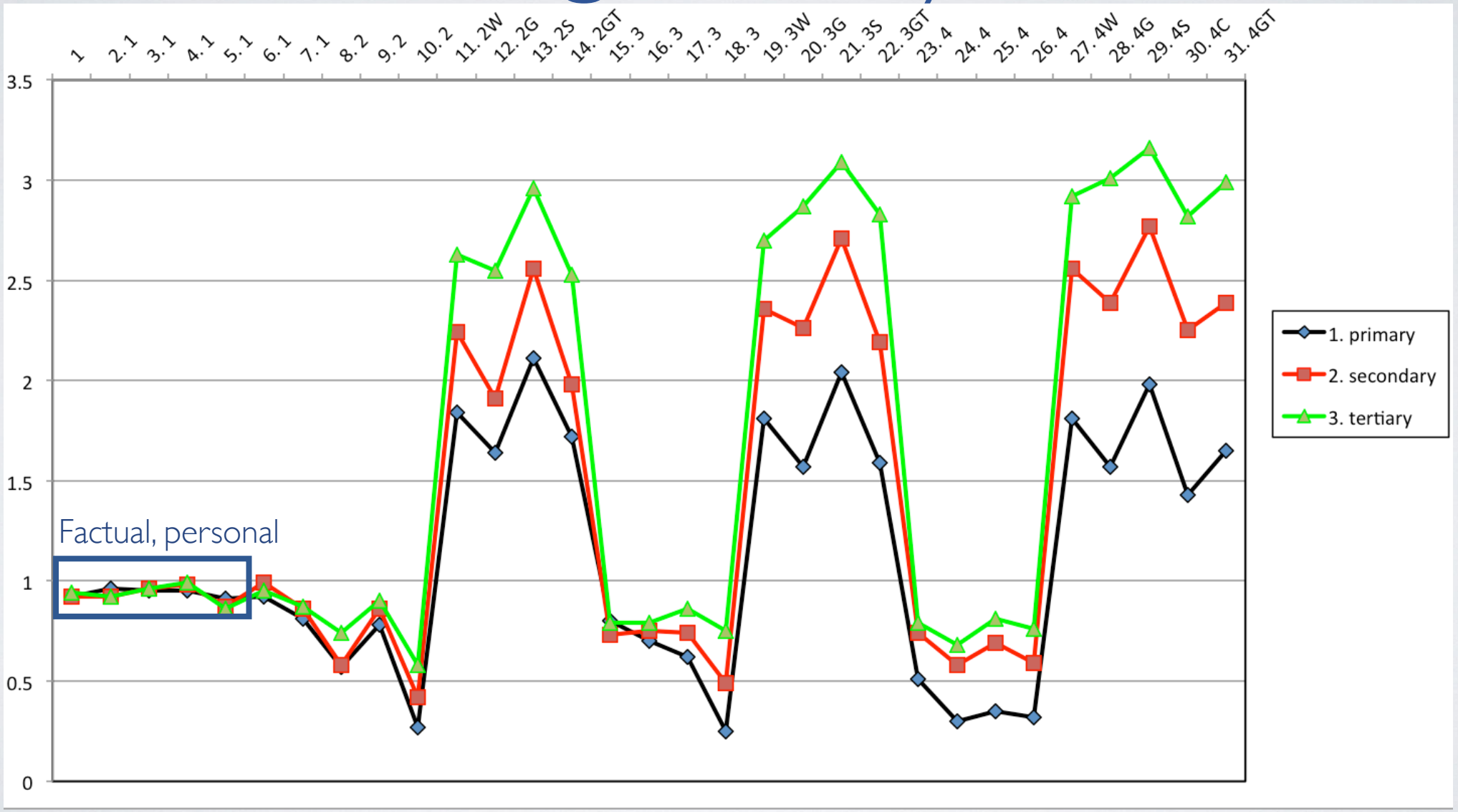


Measure	(se)	Infit	
1.83	0.04	1.23	FAST(120) A2
0.99	0.04	1.05	FAST (120) A1
0.51	0.06	0.92	FAST(160) A2
0.38	0.05	1.06	FAST(160) A1
0.23	0.3	0.95	STANDARD A2
0.10	0.3	0.88	STANDARD A1
0.07	.04	.93	SLOW(360) A2
-0.24	.06	.82	SLOW(360) A1
-0.55	.04	.80	SLOW(480) A2
-0.64	.02	.83	SLOW(480) A1
-2.17	0.23	.92	ALFA

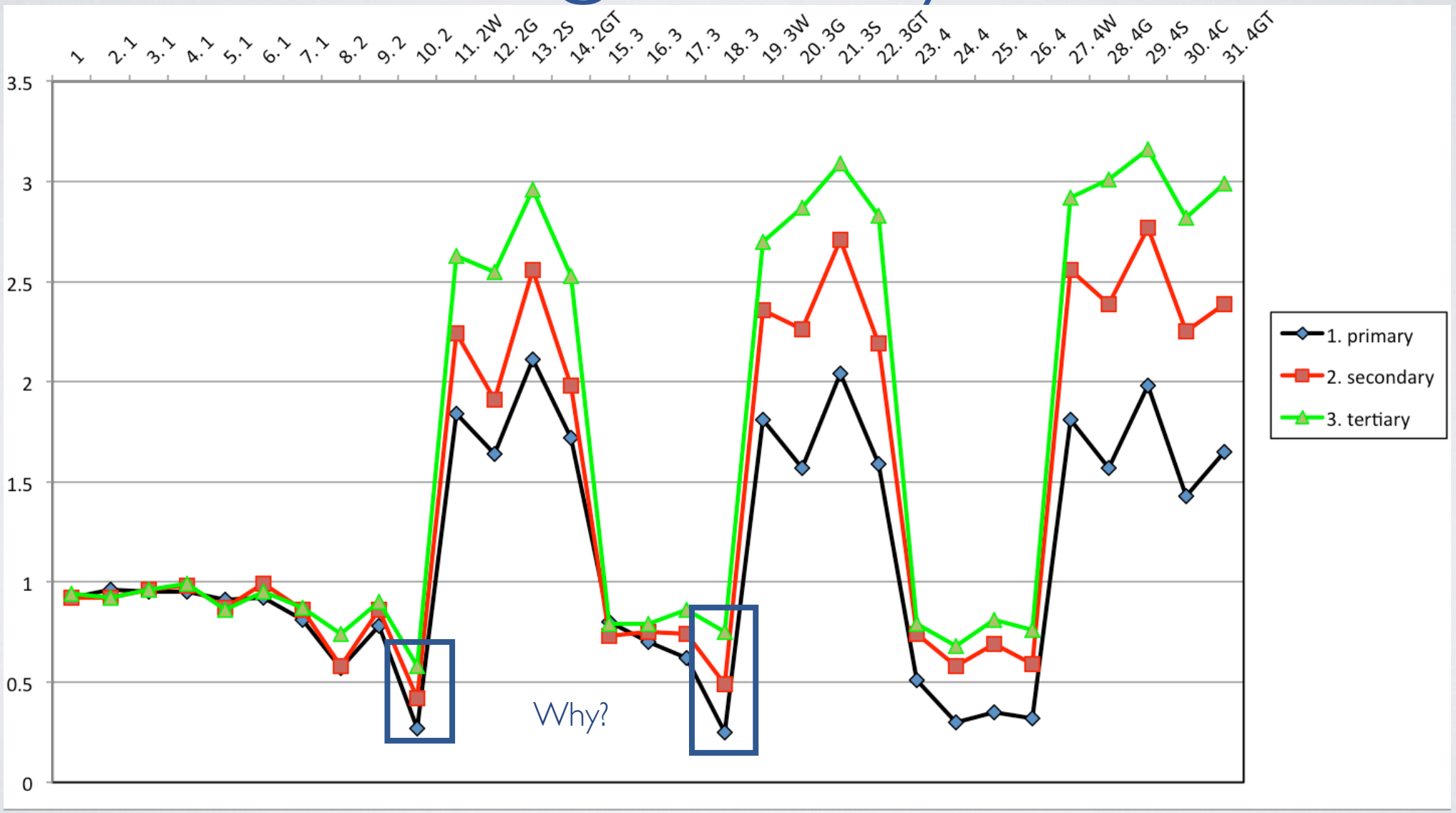
Strata 15.77 Reliability .99

$X^2(9) = 3918.7$, $p < .000$

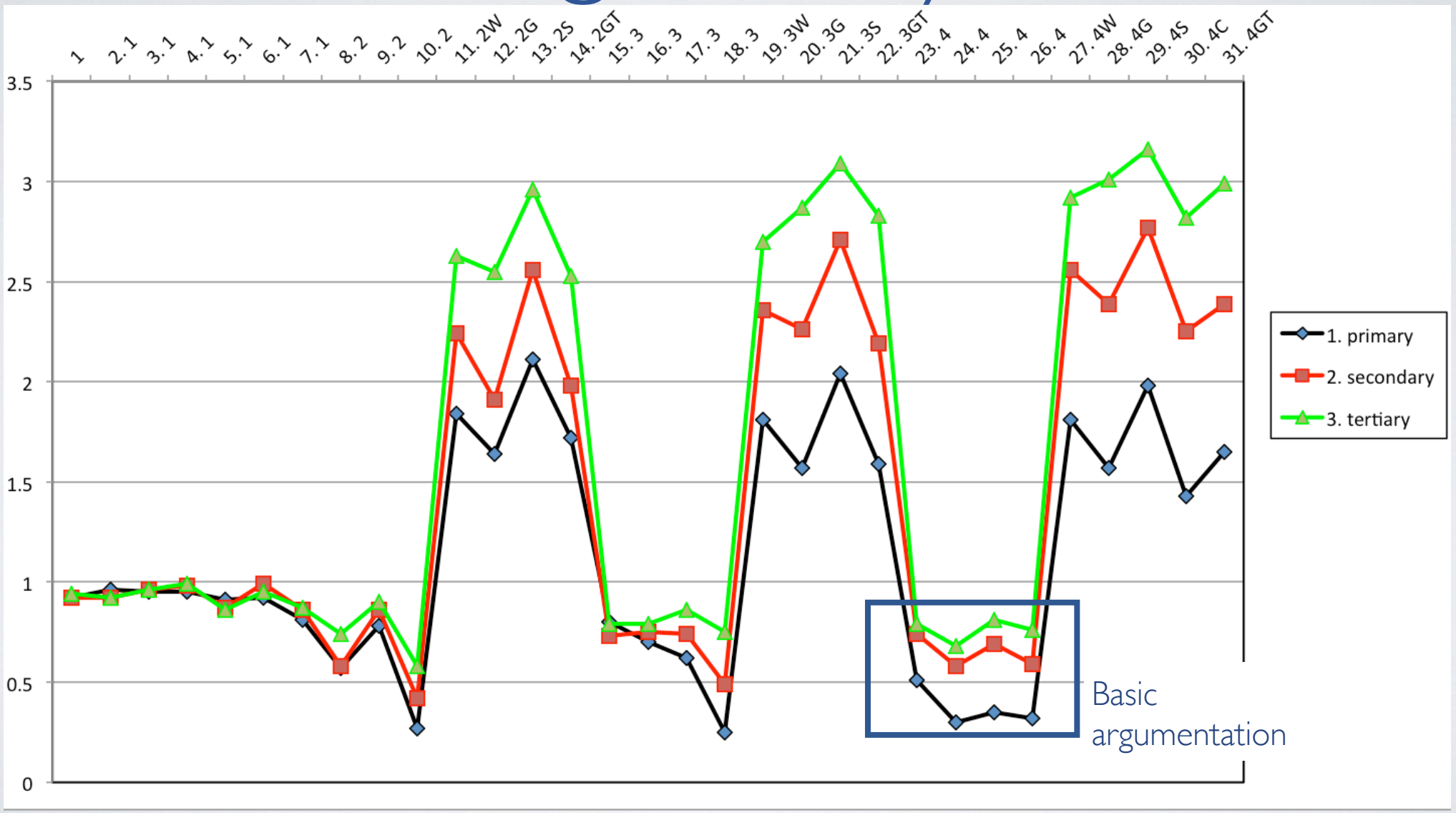
Writing: bias by item



Writing: bias by item



Writing: bias by item



Is the test fair?

Fairness = objectiveness = the absence of bias
= internal test quality

(Rawls 2001; McNamara, Knoch Fan, 2019)

Is the test fair?

Test quality
Adequate rater consistency
No item misfit or overfit

BUT

Significant performance differences primary^{-} vs primary^{+}
Ample evidence of bias

Study 1: LESLLA learners are routinely part of the general population for high-stakes tests

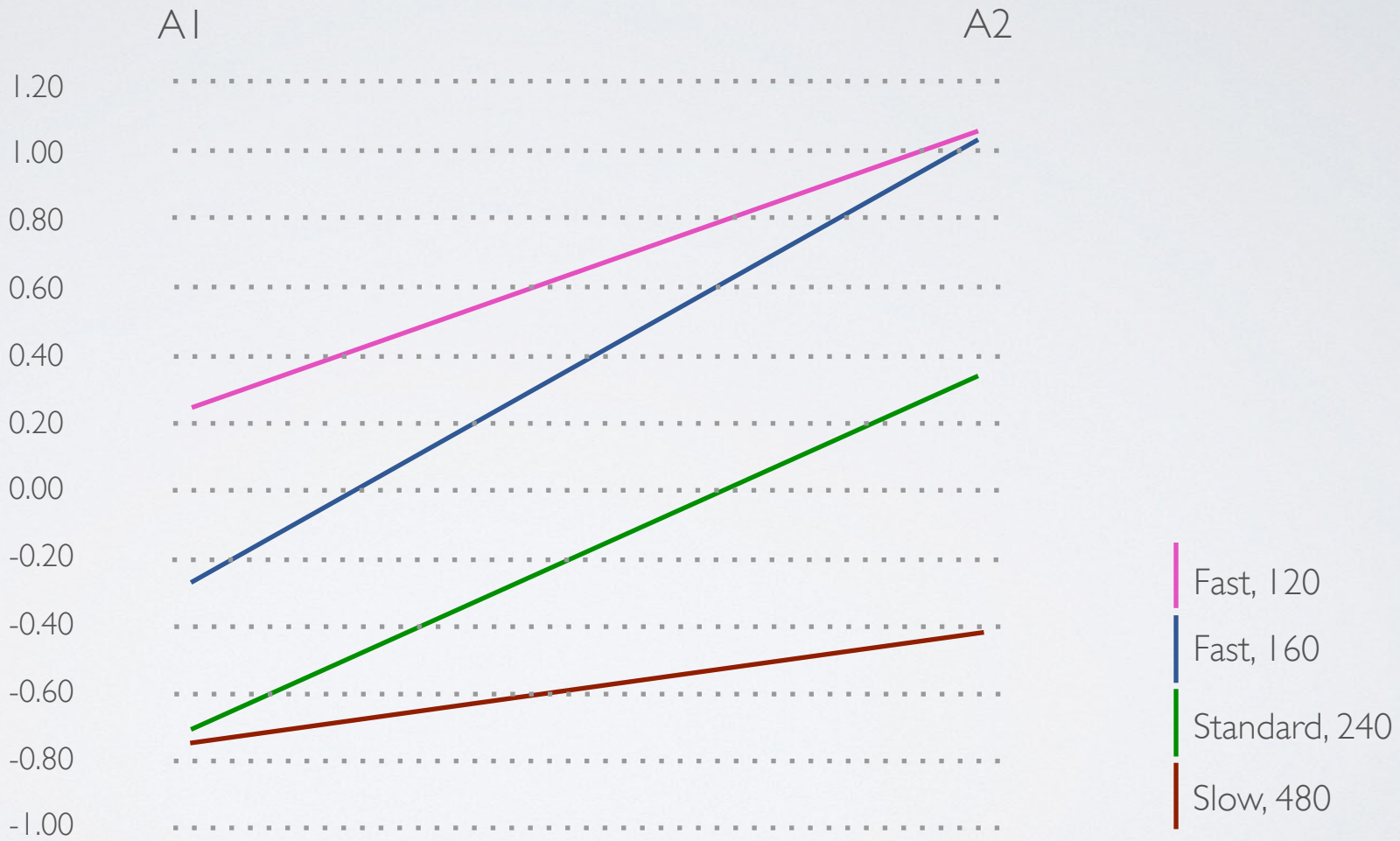
Study 2: Educational background (and track type) impact pass probability substantially and significantly

Study 1: LESLLA learners are routinely part of the general population for high-stakes tests

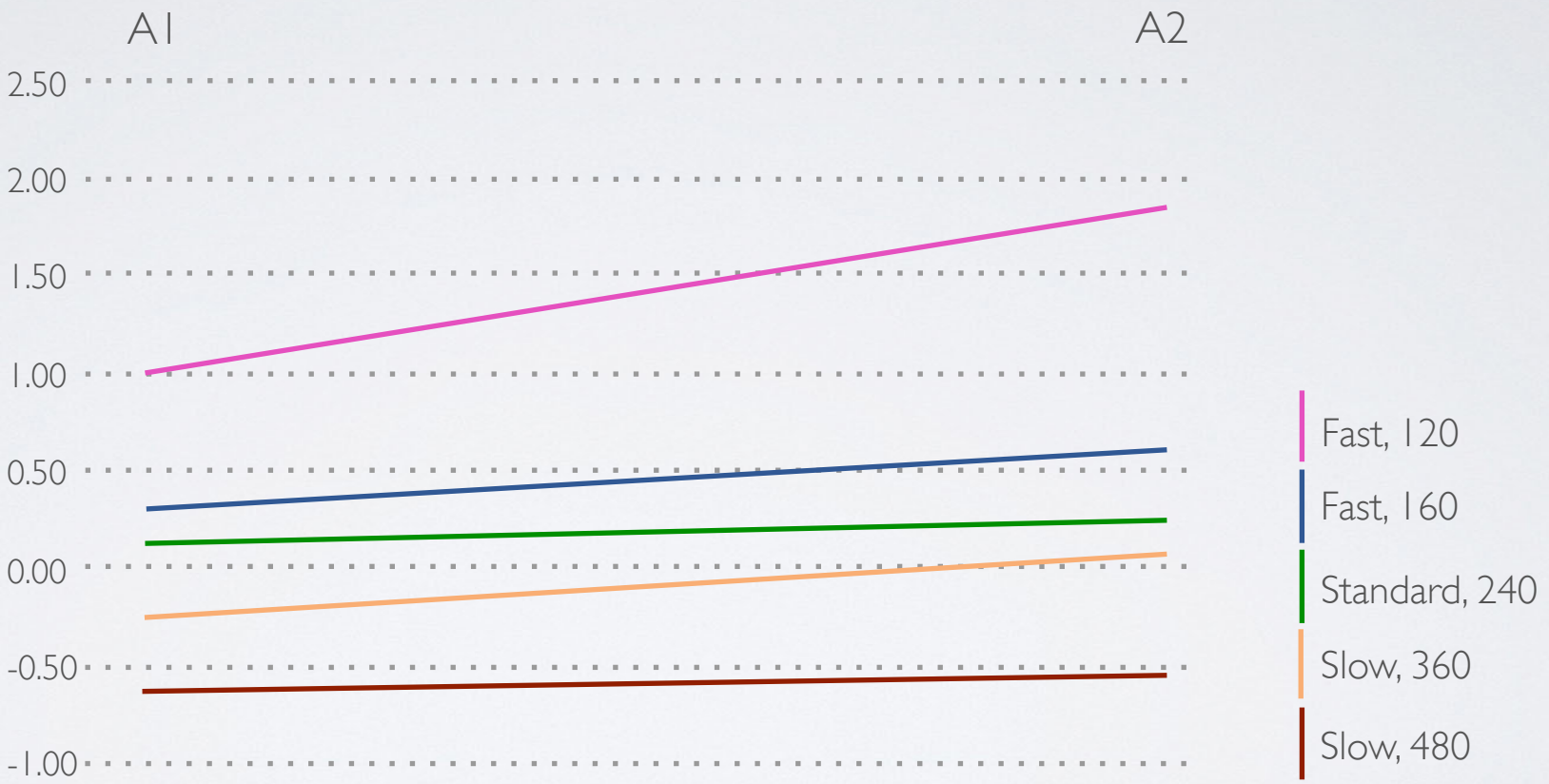
Study 2: Educational background (and track type) impact pass probability substantially and significantly

Study 3: Does teaching help?

Speaking gains (measure)



Writing gains (measure)



Writing performances in detail

N = 385

Double coded: 20% (n = 78), ICC .81 - .98

	ICC	*p*	95% CI
T-Unit	0.981	< .0000	0.965 - 0.99
Error-free TU	0.919	< .0000	0.851 - 0.957
#errors	0.949	< .0000	0.883 - 0.976
Co Clause	0.889	< .0000	0.799 - 0.94
Sub Clause	0.811	< .0000	0.623 - 0.904

Writing performances in detail

Syntactic complexity	Clauses/TU
	Words/Clause
	Mean sentence length
	Simple sentence ratio
	Compound sentence ratio
	Complex sentence ratio
	Compound complex ratio
	Coordinated clause ratio
	Subordinated clause ratio
Lexical complexity	Average word length
	Unique words / tot
	Guiraud's index
Accuracy	Incomplete sentence ratio
	Proportion of error-free T-Units
	Errors / T-Unit
	Errors / words
Fluency	Words / TU
	Total word count

(Bulté, & Housen, 2014; Iwashita., Brown., McNamara., & O'Hagan, 2008; Knoch, Rouhshad, Oon, & Storch, 2015; Serrano, Tragant, & Llanes, 2012; Treffers-Daller, Parslow,, & Williams, 2016)

Writing performance gains

Slow: No measurable gains on any indicator

Writing performance gains

Standard: Small – medium gains in accuracy

Fewer incomplete sentences ($W = 877.5, p = 0.033, d = .5$)

Less errors / T-Unit ($W = 450, p = 0.015\ d = -0.576$)

No measurable gains on syntactic / lexical complexity, fluency

Writing performance gains

Fast: Small – medium gains in syntactic complexity

Clauses / TU	($W = 1727.5$, $p = 0.002$, $d = -0.505$)
Words / Clause	($W = 3289$, $p = 0.01$ $d = 0.325$)
Simple sentence ratio	($W = 3354.5$, $p < 0.001$ $d = 0.571$)
Cx sentence ratio	($W = 1899$, $p < 0.001$ $d = -0.507$)
Subordinated clause ratio	($W = 1534.5$, $p < 0.001$ $d = -0.710$)

Writing performance gains

Fast:

Small – medium gains in syntactic complexity

- Clauses / TU ($W = 1727.5$, $p = 0.002$, $d = -0.505$)
- Words / Clause ($W = 3289$, $p = 0.01$ $d = 0.325$)
- Simple sentence ratio ($W = 3354.5$, $p < 0.001$ $d = 0.571$)
- Cx sentence ratio ($W = 1899$, $p < 0.001$ $d = -0.507$)
- Subordinated clause ratio ($W = 1534.5$, $p < 0.001$ $d = -0.710$)

Individual indicators of lexical complexity, accuracy, fluency

- Guiraud's index ($W = 2005$, $p = 0.04$, $d = -0.264$)
- Errors / total words ($W = 3104$,, $p = 0.017$, $d = -0.248$)
- Words / T-Unit ($W = 1595$,, $p < 0.000$, $d = -0.268$)

Error types, Slow vs Fast

Error type	Effect size *d*
conjunction	-0.25
article	-0.31
ellipsis	-0.34
redundancy	-0.35
spelling	-0.36
inversion	-0.4
conjugation	-0.5
non-finite clauses	-0.53
morphology	-0.55
word order	-0.57
prepositions	-0.6
pronouns	-0.62
word choice	-0.73

Does teaching even the odds?

For a test, not nearly enough
The A2 certificates are not equivalent
Bias persists

Does teaching even the odds?

But

95% of the LESLLA respondents feel welcome in Flanders

(higher educated : 92%, $W = 55550$, $p = 0.05$)

98% of the LESLLA respondents consider Flemish people friendly

(higher educated : 84%, $W = 59584$, $p < 0.000$)

75% finds a job within 2 years

Summary

Q1 Low-educated learners are an integral part of the test-taking population
Research and policy has largely ignored this group

Summary

Low-educated learners are an integral part of the test-taking population
Research and policy has largely ignored this group

Q2 Low-educated migrants in Flanders significantly and substantially underperform and have a very low pass probability

Summary

Low-educated learners are an integral part of the test-taking population
Research and policy has largely ignored this group

Low-educated migrants in Flanders significantly and substantially underperform and have a very low pass probability

Q3 Slow L2 classes deliver minimal gains, which do not level the playing field

Communication

To test developer

No hypothetical question types
Straightforward drawings
Revisit time constraints

Communication

To test developer

To language schools

More feedback
More challenging input

Communication

To test developer

To language schools

To policy makers

Think about a construct before ordering a test

Involve all stakeholders

Keep the focus on teaching

bart.deygers@kuleuven.be

Download: http://tinyurl.com/LTRCMelbourne